

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-078018
 (43)Date of publication of application : 14.03.2000

(51)Int.Cl. H03M 7/30
 G10L 19/02
 G10L 19/00

(21)Application number : 11-160383 (71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 08.06.1999 (72)Inventor : KAWAHARA EIJI

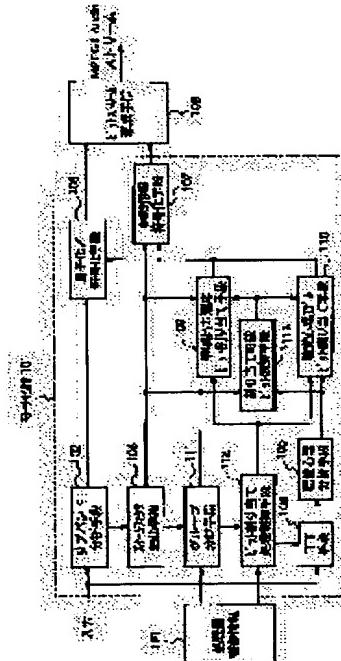
(30)Priority

Priority number : 10167245 Priority date : 15.06.1998 Priority country : JP

(54) VOICE CODING SYSTEM AND DEVICE AND DATA RECORDING MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To produce the coded data of high sound quality in real time and with no sound break by switching the using bit allocation means to execute the bit allocation and then the coding so as to perform the due processing by means of the prescribed one of plural bit allocation means.
SOLUTION: A sub-band analyzing means 102 divides the inputted digital audio signals into 32 frequency components. A grouping means 111 divides the 32 frequency components into the number of groups that is designated by the throughput control information 121. An allocatable bit arithmetic means 113 decides each allocatable bit number to an aural psychological model bit allocation means 110 and a band output adaptive bit allocation means 109 based on the total allocatable bit number. Each of sub-band signals is coded by a quantizing/coding means 106 and an auxiliary information coding means 107, and a bit stream forming means 108 forms and outputs a bit stream.



LEGAL STATUS

[Date of request for examination] 18.01.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

Best Available Copy

[Number of appeal against examiner's
decision of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japanese Patent Office

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-78018

(P2000-78018A)

(43) 公開日 平成12年3月14日 (2000.3.14)

(51) Int.Cl.⁷
 H 03 M 7/30
 G 10 L 19/02
 19/00

識別記号

F I
 H 03 M 7/30
 G 10 L 7/04
 9/18

テーマコード^{*} (参考)
 A
 G
 M

審査請求 未請求 請求項の数35 O.L. (全27頁)

(21) 出願番号 特願平11-160383
 (22) 出願日 平成11年6月8日 (1999.6.8)
 (31) 優先権主張番号 特願平10-167245
 (32) 優先日 平成10年6月15日 (1998.6.15)
 (33) 優先権主張国 日本 (JP)

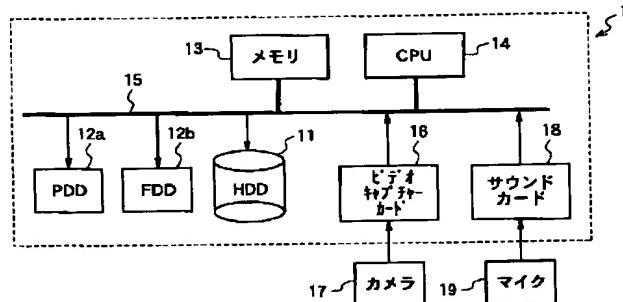
(71) 出願人 000005821
 松下電器産業株式会社
 大阪府門真市大字門真1006番地
 (72) 発明者 河原 栄治
 広島県東広島市鏡山3丁目10番18号 株式
 会社松下電器情報システム広島研究所内
 (74) 代理人 100081813
 弁理士 早瀬 嘉一

(54) 【発明の名称】 音声符号化方式、音声符号化装置、及びデータ記録媒体

(57) 【要約】

【課題】 デジタルオーディオ信号を複数の周波数帯域に分割し、各帯域ごとに符号化を行う方式において、パソコンのCPU処理能力、及び他のアプリケーションのCPU占有率に左右されず、実時間で音途切れがない高音質の符号化データの作成を実現することができる音声符号化方式を提供すること。

【解決手段】 複数の周波数帯域に分割されたデジタルオーディオ信号の各帯域に対するビット割り当て情報を生成するビット割り当て手段として、所定の聴覚心理モデルに基づく信号対マスク比値との関係を使用して高効率にビット割り当てを行う方法、及び低負荷でビット割り当てを行う方法を有し、符号化手段に対し占有できるCPUの処理量情報からビット割り当て手段を切り替え符号化を行う。



【特許請求の範囲】

【請求項 1】 デジタルオーディオ信号を複数の周波数帯域に分割し、各帯域毎に符号化を行う音声符号化方式であって、

上記分割された各帯域に対するビット割り当て情報を生成し、それぞれ処理量の異なるビット割り当て手段を複数有し、

外部からの制御情報に基づいて、上記複数のビット割り当て手段の中から、所定のビット割り当て手段を用いて処理がなされるよう、使用するビット割り当て手段を切り替えてビット割り当てを実行して、符号化を行うことを特徴とする音声符号化方式。

【請求項 2】 請求項 1 記載の音声符号化方式において、

上記外部からの制御情報として、符号化処理を行うために占有できる中央演算処理装置の処理量を表す負荷値を用い、

上記負荷値に基づいて、上記中央演算処理装置上で符号化処理において各ビット割り当て手段を用いて符号化動作を行った時の各処理量を予め記憶したデータテーブルを参照して、上記符号化処理が占有できる中央演算処理装置の処理量を超えないよう上記ビット割り当て手段の選択を行うことを特徴とする音声符号化方式。

【請求項 3】 請求項 2 記載の音声符号化方式において、

上記負荷値として、符号化処理を行うために占有可能な上記中央演算処理装置の処理量を監視する監視手段からの処理量制御情報を用いることを特徴とする音声符号化方式。

【請求項 4】 請求項 1 記載の音声符号化方式において、
上記ビット割り当て手段によるビット割り当て処理として、

符号化データの高音質化を実現可能な高効率ビット割り当てを行う高効率ビット割り当て方法を用いた処理と、

上記高効率ビット割り当て方法を用いた処理と比較して処理量の少ない低負荷でビット割り当てを行う低負荷ビット割り当て方法を用いた処理とが行なわれることを特徴とする音声符号化方式。

【請求項 5】 請求項 1 記載の音声符号化方式において、

上記符号化時に使用されるビット割り当て手段の切り替えを、オーディオ信号に復号可能な最小単位であるフレーム単位で行うことを特徴とする音声符号化方式。

【請求項 6】 請求項 1 記載の音声符号化方式において、

複数の周波数帯域に分割された各帯域のサブバンド信号を、各々予め定められた所定個数のサブバンド信号からなるグループとなるようにグループ分けを行い、各グル

ープに対して独立したビット割り当て処理を行い、各帯域に対するビット割り当て情報を生成することを特徴とする音声符号化方式。

【請求項 7】 請求項 6 記載の音声符号化方式において、

上記グループ分けは、

グループの数、又はグループ内の周波数軸方向に連続したサブバンド信号の数を、上記外部からの制御情報により指定された数、又は上記監視手段からの処理量制御情報に基づいて指定された数になるよう、可変的に行われるなどを特徴とする音声符号化方式。

【請求項 8】 請求項 7 記載の音声符号化方式において、

上記サブバンド信号の数の変更処理を、オーディオ信号に復号可能な最小単位であるフレーム単位で行うこととする音声符号化方式。

【請求項 9】 請求項 8 記載の音声符号化方式において、

上記グループ分け時に、ビット割り当てを行わないグループを少なくとも 1 つ設けることを特徴とする音声符号化方法。

【請求項 10】 請求項 6 記載の音声符号化方式において、

上記サブバンド信号のグループ分けにより、低帯域に属するグループにグループ分けされたサブバンド信号に対し、符号化データの高音質化を実現可能な高効率ビット割り当てを行う処理を用いて、グループ内のサブバンド信号にビット割り当てを行い、一方、高帯域に属するグループにグループ分けされたサブバンド信号に対し、上記高効率ビット割り当て処理と比較して処理量の少ない低負荷ビット割り当て処理を用いて、グループ内のサブバンド信号にビット割り当てを行うことを特徴とする音声符号化方式。

【請求項 11】 請求項 6 記載の音声符号化方式において、

各グループ毎に独立したビット割り当て手段に対する割り当て可能ビット数を決定する割り当て可能ビット演算手段を設け、各グループのグループ全体に対する割合に、各グループに属する各帯域毎の特性に基づいた重み付けを加味したものを用いて、グループ全体に対する割り当て可能ビット数を、各グループ毎に独立したビット割り当て手段に対し振り分けることを特徴とする音声符号化方式。

【請求項 12】 請求項 11 記載の音声符号化方式において、

各グループに属する各帯域毎の特性に基づいた重み付けを、各帯域毎の所定の最小可聴限界値に基づいた重み付けとすることを特徴とする音声符号化方式。

【請求項 13】 請求項 11 記載の音声符号化方式において、

各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号にサブバンド分析を施して得られる、各グループに属する各周波数帯域のサブバンド信号レベルに基づいた重み付けとすることを特徴とする音声符号化方式。

【請求項 14】 請求項 11 記載の音声符号化方式において、

各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号に線形変換を施して得られる、各グループに属するスペクトル信号レベルに基づいた重み付けとすることを特徴とする音声符号化方式。

【請求項 15】 請求項 6 記載の音声符号化方式において、

各グループに属する信号レベルが、所定のしきい値以上の高レベルな信号に対しては、符号化データの高音質化を実現可能な高効率ビット割り当て方法を用いてビット割り当てを行い、

各グループに属する信号レベルが、所定のしきい値以下の低レベルな信号に対しては、上記高効率ビット割り当て方法と比較して処理量の少ない低負荷ビット割り当て方法を用いてビット割り当てを行うことを特徴とする音声符号化方式。

【請求項 16】 請求項 15 記載の音声符号化方式において、

上記各グループに属する信号レベルを、入力デジタルオーディオ信号にサブバンド分析を施して得られるサブバンド信号レベルとすることを特徴とする音声符号化方式。

【請求項 17】 請求項 15 記載の音声符号化方式において、

上記各グループに属する信号レベルが、入力デジタルオーディオ信号に線形変換を施して得られるスペクトル信号レベルであることを特徴とする音声符号化方式。

【請求項 18】 請求項 15 記載の音声符号化方式において、

上記各グループに属する信号レベルを、所定の各帯域毎の最小可聴限界値とすることを特徴とする音声符号化方式。

【請求項 19】 請求項 4, 10, 15 のいずれかに記載の音声符号化方式において、

上記符号化データの高音質化を実現可能な高効率ビット割り当て処理は、所定の聴覚心理モデルに基づく信号対マスク比値との関係を使用して行われるビット割り当て処理であり、

上記高効率ビット割り当て処理と比較して処理量の少ない低負荷ビット割り当て処理は、複数の周波数帯域に分割された信号レベルに各帯域毎の所定の最小可聴限界値を加味して行われるビット割り当て処理であることを特徴とする音声符号化方式。

【請求項 20】 請求項 19 記載の音声符号化方式において、

上記聴覚心理モデルが、MPEG (Motion Picture Experts Group) によって指定された聴覚心理モデルであることを特徴とする音声符号化方式。

【請求項 21】 請求項 5 または請求項 8 記載の音声符号化方式において、

上記オーディオ信号に復号可能な最小単位であるフレームが、MPEG (Motion Picture Experts Group) によって指定されたフレームであることを特徴とする音声符号化方式。

【請求項 22】 請求項 1 記載の音声符号化方式において、

上記ビット割り当て手段は、分割化された各帯域に対し、所定の聴覚心理モデルから出力される情報に基づいてビット割り当て情報を生成するものであり、 $N(N=1, 2, 3\dots)$ フレームに 1 度、上記所定の聴覚心理モデルから出力される情報に基づいてビット割り当て情報を生成し、

上記ビット割り当て情報を生成しなかったフレームに対しては、上記聴覚心理モデルから出力された情報と上記分割された各帯域の信号情報に基づいてビット割り当て情報を生成し、符号化を行なうことを特徴とする音声符号化方式。

【請求項 23】 請求項 1 記載の音声符号化方式において、

段階的に処理量の制御可能な聴覚心理モデルを有し、外部からの制御情報に基づいて、上記聴覚心理モデルの処理量制御を行ない、所定の処理量の聴覚心理モデルを用いて処理がなされるように、各帯域に対するビット割り当て情報を生成することを特徴とする音声符号化方式。

【請求項 24】 請求項 1 記載の音声符号化方式において、

それぞれ処理量の異なる聴覚心理モデルを複数有し、外部からの制御情報に基づいて、上記複数の聴覚心理モデルの中から、所定の聴覚心理モデルを用いて処理が成されるように、使用する聴覚心理モデルを切り替えて、各帯域に対するビット割り当て情報を生成することを特徴とする音声符号化方式。

【請求項 25】 デジタルオーディオ信号を複数の周波数帯域に分割し、該分割された各帯域に対するビット割り当て情報を生成し、所定ビットレートでの伝送を目的として各帯域毎に符号化を行なう音声符号化方式であつて、

符号化データストリーム中にデータを挿入するフレームにおけるビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御することを特徴とする音声符号化方式。

【請求項 26】 請求項 25 記載の音声符号化方式にお

いて、

外部からの制御情報に基づいて、フレーム単位でビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御することを特徴とする音声符号化方式。

【請求項27】 請求項26記載の音声符号化方式において、

上記外部からの制御情報として、付加データのバッファを監視する監視手段からのデータ量制御情報を用いることを特徴とする音声符号化方式。

【請求項28】 請求項1記載の音声符号化方式において、

符号化処理が実行される中央演算処理装置の性能に応じて、符号化処理動作前の初期化時に複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報を外部へ出力することを特徴とする音声符号化方式。

【請求項29】 請求項28記載の音声符号化方式において、

外部への情報として出力される複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報が、降順、あるいは昇順で出力されることを特徴とする音声符号化方式。

【請求項30】 映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、

異なる複数の演算量で符号化を行うものであって、

上記音声信号もしくは映像信号の符号化の演算量を変更することで、上記中央演算処理装置上での処理の全体的な演算量を制御することを特徴とする音声符号化方式。

【請求項31】 映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、

演算量の異なる複数の符号化方式を用いて符号化を行うものであって、

上記音声信号の符号化の符号化方式を変更することで、上記中央演算処理装置上での処理の全体的な演算量を制御することを特徴とする音声符号化方式。

【請求項32】 請求項30または31記載の音声符号化方式において、

上記中央演算処理装置上での処理の制御を、外部からの制御情報に基づいて行うことを特徴とする音声符号化方式。

【請求項33】 デジタルオーディオ信号に対して、時間／周波数変換を施し、量子化情報を生成することで符号化を行う音声符号化方式であって、

それぞれ演算量の異なる複数の量子化情報算出手段を有し、

外部からの制御情報に基づいて、上記複数の量子化情報算出手段の中から、所定の量子化情報算出手段を選択して処理が行われるように、使用する量子化情報算出手段を切り替えて量子化情報を算出して符号化を行うことを

特徴とする音声符号化方式。

【請求項34】 請求項1ないし請求項33のいずれかに記載の音声符号化方式を用いて音声符号化を行うことを特徴とする音声符号化装置。

【請求項35】 請求項1ないし請求項33のいずれかに記載の音声符号化方式のステップが記録されていることを特徴とするデータ記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声符号化方式、及び音声符号化装置、及びデータ記憶媒体に関し、特にMPEG (Motion Picture Experts Group) 方式で用いられているようなサブバンド符号化方式を用いる音声符号化方式、及び音声符号化装置、及び前記音声符号化方式を実行するためのプログラムを格納したデータ記憶媒体に関するものである。

【0002】

【従来の技術】近年、パソコンのマルチメディア化やインターネットの普及により、パソコン（以下、PCともいう）等の上で、ソフトウェアによってMPEG等の動画や音声を再生できる環境が整ってきており、MPEG等の符号化データの利用範囲が広がっている。しかしながら、符号化データを作るエンコーダに関しては、いまだに高価なハードウェアを用いられるのが主流である。また、ソフトウェアで符号化データを作るものもあるが、符号化対象となる動画や音声の再生時間の実時間の何倍もの処理時間をかけて符号化を行うものであるため、多大な時間、及び手間を要し、広く普及するに至っていない。

【0003】このため、特に一般的のパソコンユーザが安価で簡単に符号化データを作成できるようになるためにには、ソフトウェア処理により実時間での符号化データの作成を実現したいという要望がある。

【0004】以下に従来の音声符号化方式の一例について説明する。図11は音声に関する符号化データフォーマットとして、ISO/IEC11172-3にて規格化されているMPEGオーディオエンコーダのブロック図である。図11において、入力デジタルオーディオ信号は、サブバンド分析手段202において32個の周波数成分に分割され、各サブバンド信号に対し、スケールファクタ抽出手段203においてスケールファクタを計算し、ダイナミックレンジをそろえる。また、入力デジタルオーディオ信号は、FFT手段204において、高速フーリエ変換(FFT:Fast Fourier Transform)され、この結果を用いて、聴覚心理分析手段205により人間の聴覚の特性を利用した聴覚心理モデルに基づく信号対マスク比(SMR)値の関係モデルを利用し、ビット割り当て手段206により各サブバンド信号に対するビット割り当て数を決める。各サブバンド信号へのビット割り当て数に応じて、各サブバンド信号を量子化・符号化手段207によ

り量子化／符号化する。そして、ビットストリーム形成手段209により、補助情報符号化手段208により符号化されたヘッダ情報と補助情報を共にしてビットストリームを形成して出力する。

【0005】この従来の音声符号化方式は、各帯域電力の偏在を利用して各帯域（サブバンド）毎に符号化を行うような符号化方式であるため、聴覚心理モデルを利用した各サブバンド信号に対するピット配分が音質を左右することになる。また、蓄積媒体を利用して規格化されたため、符号化データの高音質化には適しているが、実時間での符号化には適しておらず、音質を左右する聴覚心理モデルは非常に演算量の多いものとなっている。

【0006】

【発明が解決しようとする課題】従来の音声符号化方式、及び音声符号化装置は以上のように構成されており、蓄積媒体を対象とする高音質な符号化データを作成するには適しているが、聴覚心理モデルの利用は多くの処理能力が必要とされるため、ソフトウエアで処理するには現在のCPU能力ではPC上で実時間処理するには不適当である。また、実時間処理可能な高性能なCPUを搭載したPC上で動作させた場合においても、他のアプリケーションのCPU占有率が大きくなつた時などでは、実時間での処理が不可能となる恐れがあり、その結果、音途切れが発生する可能性があるという問題点がある。

【0007】この発明は、以上のような問題点を解決するためになされたもので、パソコンのCPU処理能力、及び他のアプリケーションのCPU占有率に左右されず、実時間で音途切れない高音質の符号化データの作成をソフトウエア処理により実現することができる音声符号化方式、音声符号化装置、及び上記符号化を実行するためのプログラムを格納したデータ記憶媒体を提供することを目的とする。

【0008】

【課題を解決するための手段】この発明の請求項1にかかる音声符号化方式は、デジタルオーディオ信号を複数の周波数帯域に分割し、各帯域毎に符号化を行う音声符号化方式であつて、上記分割された各帯域に対するピット割り当て情報を生成し、それぞれ処理量の異なるピット割り当て手段を複数有し、外部からの制御情報に基づいて、上記複数のピット割り当て手段の中から、所定のピット割り当て手段を用いて処理がなされるように、使用するピット割り当て手段を切り替えてピット割り当てを実行して、符号化するものである。

【0009】また、本発明の請求項2にかかる音声符号化方式は、上記請求項1記載の音声符号化方式において、上記外部からの制御情報として、符号化処理を行るために占有できる中央演算処理装置の処理量を表す負荷値を用い、該負荷値に基づいて上記中央演算処理装置で

符号化処理において各ピット割り当て手段を用いて符号化動作を行つた時の各処理量を予め記憶したデータテーブルを参照して、上記符号化処理が占有できる中央演算処理装置の処理量を超えないよう上記ピット割り当て手段の選択を行うものである。

【0010】また、本発明の請求項3にかかる音声符号化方式は、上記請求項2記載の音声符号化方式において、上記負荷値として、符号化処理を行うために占有可能な上記中央演算処理装置の処理量を監視する監視手段からの処理量制御情報を用いるものである。

【0011】また、本発明の請求項4にかかる音声符号化方式は、上記請求項1記載の音声符号化方式において、上記ピット割り当て手段によるピット割り当て処理として、符号化データの高音質化を実現可能な高効率にピット割り当てを行う高効率ピット割り当て方法を用いた処理と、上記高効率ピット割り当て方法を用いた処理と比較して処理量の少ない低負荷でピット割り当てを行う低負荷ピット割り当て方法を用いた処理とが行われるものである。

【0012】また、本発明の請求項5にかかる音声符号化方式は、上記請求項1記載の音声符号化方式において、上記符号化時に使用されるピット割り当て手段の切り替えを、オーディオ信号に復号可能な最小単位であるフレーム単位で行うものである。

【0013】また、本発明の請求項6にかかる音声符号化方式は、上記請求項1記載の音声符号化方式において、複数の周波数帯域に分割された各帯域のサブバンド信号を、各々予め定められた所定個数のサブバンド信号からなるグループとなるように分けを行うグループ分けを行い、各グループに対して独立したピット割り当て処理を行い、各帯域に対するピット割り当て情報を生成するものである。

【0014】また、本発明の請求項7にかかる音声符号化方式は、上記請求項6記載の音声符号化方式において、上記グループ分けは、グループの数、又はグループ内の周波数軸方向に連続したサブバンド信号の数を、上記外部からの制御情報により指定された数、又は上記監視手段からの処理量制御情報に基づいて指定された数に可変的に行われるものである。

【0015】また、本発明の請求項8にかかる音声符号化方式は、上記請求項7記載の音声符号化方式において、上記サブバンド信号の数の変更処理を、オーディオ信号に復号可能な最小単位であるフレーム単位で行うものである。

【0016】また、本発明の請求項9にかかる音声符号化方式は、上記請求項8記載の音声符号化方式において、上記グループ分け時に、ピット割り当てを行わないグループを少なくとも1つ設けるものである。

【0017】また、本発明の請求項10にかかる音声符号化方式は、上記請求項6記載の音声符号化方式において、

て、上記サブバンド信号のグループ分けにより、低帯域に属するグループにグループ分けされたサブバンド信号に対し、符号化データの高音質化を実現可能な高効率ビット割り当て方法を用いて、グループ内のサブバンド信号にビット割り当てを行い、グループ内のサブバンド信号にビット割り当てを行つ一方、高帯域に属するグループにグループ分けされたサブバンド信号に対し、高効率ビット割り当て方法と比較して処理量の少ない低負荷ビット割り当て処理を用いて、グループ内のサブバンド信号にビット割り当てを行つものである。また、本発明の請求項11にかかる音声符号化方式は、上記請求項6記載の音声符号化方式において、各グループ毎に独立したビット割り当て手段に対する割り当て可能ビット数を決定する割り当て可能ビット演算手段を設け、各グループのグループ全体に対する割合に各グループに属する各帯域毎の特性に基づいた重み付けを加味したもの用いて、グループ全体に対する割り当て可能ビット数を各グループ毎に独立したビット割り当て手段に対し振り分けるものである。

【0018】また、本発明の請求項12にかかる音声符号化方式は、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、各帯域毎の所定の最小可聴限界値に基づいた重み付けとするものである。また、本発明の請求項13にかかる音声符号化方式は、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号にサブバンド分析を施して得られる各グループに属する各周波数帯域のサブバンド信号レベルに基づいた重み付けとするものである。

【0019】また、本発明の請求項14にかかる音声符号化方式は、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号に線形変換を施して得られる各グループに属するスペクトル信号レベルに基づいた重み付けとするものである。

【0020】また、本発明の請求項15にかかる音声符号化方式は、上記請求項6記載の音声符号化方式において、各グループに属する信号レベルが所定のしきい値以上の高レベルな信号に対しては、符号化データの高音質化を実現可能な高効率ビット割り当て方法を用いてビット割り当てを行い、各グループに属する信号レベルが、所定のしきい値以下の低レベルな信号に対しては、上記高効率ビット割り当て方法と比較して処理量の少ない低負荷ビット割り当て方法を用いてビット割り当てを行つものである。

【0021】また、本発明の請求項16にかかる音声符号化方式は、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルを、入力デジタルオーディオ信号にサブバンド分析を施して得られる

サブバンド信号レベルとするものである。

【0022】また、本発明の請求項17にかかる音声符号化方式は、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルが、入力デジタルオーディオ信号に線形変換を施して得られるスペクトル信号レベルであるものである。

【0023】また、本発明の請求項18にかかる音声符号化方式は、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルを、所定の各帯域毎の最小可聴限界値とするものである。

【0024】また、本発明の請求項19にかかる音声符号化方式は、上記請求項4、10、15のいずれかに記載の音声符号化方式において、上記符号化データの高音質化を実現可能な高効率ビット割り当て処理は、所定の聴覚心理モデルに基づく信号対マスク比値との関係を使用して行われるビット割り当て処理であり、上記高効率ビット割り当て処理と比較して処理量の少ない低負荷ビット割り当て処理は、複数の周波数帯域に分割された信号レベルに各帯域毎の所定の最小可聴限界値を加味して行われるビット割り当て処理であるものとしたものである。

【0025】また、本発明の請求項20にかかる音声符号化方式は、上記請求項19記載の音声符号化方式において、聴覚心理モデルがMPEG (Motion Picture Experts Group) によって指定された聴覚心理モデルであるものである。

【0026】また、本発明の請求項21記載の音声符号化方式は、上記請求項5または請求項8記載の音声符号化方式において、上記オーディオ信号に復号可能な最小単位であるフレームが、MPEG (Motion Picture Experts Group) によって指定されたフレームであるものである。

【0027】また、本発明の請求項22記載の音声符号化方式は、上記請求項1記載の音声符号化方式において、上記ビット割り当て手段は、分割化された各帯域に對し、所定の聴覚心理モデルから出力される情報に基づいてビット割り当て情報を生成するものであり、 $N(N=1, 2, 3\dots)$ フレームに1度、上記所定の聴覚心理モデルから出力される情報に基づいてビット割り当て情報を生成し、上記ビット割り当て情報を生成しなかつたフレームに対しては、上記聴覚心理モデルから出力された情報と上記分割された各帯域の信号情報に基づいてビット割り当て情報を生成し、符号化を行なうものである。

【0028】また、本発明の請求項23記載の音声符号化方式は、上記請求項1記載の音声符号化方式において、段階的に処理量の制御可能な聴覚心理モデルを有し、外部からの制御情報に基づいて、上記聴覚心理モデルの処理量制御を行ない、所定の処理量の聴覚心理モデルを用いて処理がなされるように、各帯域に対するビット割り当て情報を生成するものである。

【0029】また、本発明の請求項24記載の音声符号化方式は、上記請求項1記載の音声符号化方式において、それぞれ処理量の異なる聴覚心理モデルを複数有し、外部からの制御情報に基づいて、上記複数の聴覚心理モデルの中から、所定の聴覚心理モデルを用いて処理が成されるように、使用する聴覚心理モデルを切り替えて、各帯域に対するビット割り当て情報を生成するものである。

【0030】また、本発明の請求項25記載の音声符号化方式は、デジタルオーディオ信号を複数の周波数帯域に分割し、該分割された各帯域に対するビット割り当て情報を生成し、所定ビットレートでの伝送を目的として各帯域毎に符号化を行なう音声符号化方式であって、符号化データストリーム中にデータを挿入するフレームにおけるビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御するものである。

【0031】また、本発明の請求項26記載の音声符号化方式は、上記請求項25記載の音声符号化方式において、外部からの制御情報に基づいて、フレーム単位でビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御するものである。

【0032】また、本発明の請求項27にかかる音声符号化方式は、上記請求項26記載の音声符号化方式において、上記外部からの制御情報として、付加データのバッファを監視する監視手段からのデータ量制御情報を用いるものである。

【0033】また、本発明の請求項28にかかる音声符号化方式は、上記請求項1記載の音声符号化方式において、符号化処理が実行される中央演算処理装置の性能に応じて、符号化処理動作前の初期化時に複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報を外部へ出力するものである。

【0034】また、本発明の請求項29にかかる音声符号化方式は、上記請求項28記載の音声符号化方式において、外部への情報として出力される複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報が、降順、あるいは昇順で出力されるものである。

【0035】また、本発明の請求項30にかかる音声符号化方式は、映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、異なる複数の演算量で符号化を行うものであつて、上記音声信号もしくは映像信号の符号化の演算量を変更することで、上記中央演算処理装置上での処理の全体的な演算量を制御するものである。

【0036】また、本発明の請求項31にかかる音声符号化方式は、映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、演算量の異なる複数の符号化方式を用いて符号化を行うものであつて、上記音声信号の符号化の符号化方式を変更することで、上記中央演算処理装置上での処

理の全体的な演算量を制御するものである。

【0037】また、本発明の請求項32にかかる音声符号化方式は、上記請求項30または31記載の音声符号化方式において、上記中央演算処理装置上での処理の制御を、外部からの制御情報に基づいて行うものである。

【0038】また、本発明の請求項33にかかる音声符号化方式は、デジタルオーディオ信号に対して、時間／周波数変換を施し、量子化情報を生成することで符号化を行う音声符号化方式であつて、それぞれ演算量の異なる複数の量子化情報算出手段を有し、外部からの制御情報に基づいて、上記複数の量子化情報算出手段の中から、所定の量子化情報算出手段を選択して処理が行われるように、使用する量子化情報算出手段を切り替えて量子化情報を算出して符号化を行うものである。

【0039】また、本発明の請求項34記載の音声符号化装置は、上記請求項1ないし請求項33のいずれかに記載の音声符号化方式を用いて音声符号化を行うものである。

【0040】また、本発明の請求項35記載の記録媒体は、上記請求項1ないし請求項33のいずれかに記載の音声符号化方式のステップが記録されているものである。

【0041】

【発明の実施の形態】以下、本発明の実施の形態による音声符号化方式、音声符号化装置について、図面を参照しながら説明する。

(実施の形態1) ここでは、例えば、入力信号を複数の周波数成分に分割し、各帯域電力の偏在を利用して各帯域(サブバンド)毎に符号化を行うような符号化方式の場合について説明する。図1は例えれば、上記符号化方式による音声符号化装置としてパーソナルコンピュータ

(以下、「PC」ともいう)が用いられた場合のシステム全体の概念図であり、図において、1はカメラ17とマイク19などの外部機器よりのデータ入力が可能な、いわゆるマルチメディア型のPCであり、各種データ及びプログラムを格納する記憶容量が大きく固定式の記録媒体であるハードディスクドライブ(HDD)11と、HDD11とプログラムやデータなどの入出力を行うための比較的記憶容量の小さな着脱自在な記憶媒体であるPDドライブ12a、FDドライブ12bを有し、上記HDDに格納されたプログラムが中央演算処理装置(CPU)14からの命令により適宜ランダムアクセスメモリ(RAM)などで構成されたメモリ13上に読み出されて実行されるように構成されている。また、外部機器であるカメラ17、マイク19の映像、音声を取り込むために、それぞれビデオキャプチャーカード16、サウンドカード18が内蔵されている。そして以上のような構成を有するPC1は、内部のデータバス15によって各要素が接続されている。

【0042】図2は図1に示したPC1により実行され

る音声符号化処理を実現するための音声符号化装置の符号化器 20 のブロック構成図であり、実際には H D 1 1 からメモリ 13 に読み出されたプログラムによって実現されているものである。図 2において、21 は C P U 1 4 の負荷状態を監視するための C P U 負荷監視情報であり、22 は C P U 負荷監視情報 21 に基づいて、低帯域符号化処理手段 23 と高帯域符号化処理手段 24 の動作を制御する符号化手段制御手段である。また、25 は上記 2つの符号化手段 23、24 の出力をそれぞれストリーム信号とするためのビットストリーム形成処理手段である。さらに、26 は、ユーザが指定することにより上記符号化手段制御手段 22 に入力される符号化モード指定信号である。

【0043】図 2における低帯域符号化処理手段 23 の構成としては、例えば、図 1 1 の従来例に示したような構成を用いられる。また、高帯域符号化処理手段 24 の構成としては、例えば、図 3に示すように、図 1 1 の例と同様に、各帯域電力の偏在を利用して各帯域（サブバンド）毎に符号化を行うような符号化方式を用いているが、聴覚心理モデルを利用した各サブバンド信号に対するビット配分は行なわず、その代替手段として、帯域出力適応ビット割り当て手段 304 を設け、サブバンド信号毎のスケールファクタに人間の聴覚特性に基づいた重み付けを行い、高音質化よりも低負荷処理を第 1 の目的とし、演算量の少ない構成とする。また、特定帯域への過度なビット割り当て集中を排除するために、ビットの割り当て毎に各帯域に応じた重み付けの調整を行なうものとしている。

【0044】図 4 は図 2 に示した符号化器 20 の詳細な構成を示すブロック図であり、101 は符号化器であり、後述するサブバンド分析手段 102、スケールファクタ抽出手段 103、F F T 手段 104、聴覚心理分析手段 105、量子化／符号化手段 106、補助情報符号化手段 107、ビットストリーム形成手段 108、帯域出力適応ビット割り当て手段 109、聴覚心理モデルビット割り当て手段 110、グループ分け手段 111、ビット割り当て処理制御手段 112、割り当て可能ビット演算手段 113 から構成されている。

【0045】上記サブバンド解析手段 102 は、入力されたデジタルオーディオ信号を 32 個の周波数成分に分割する。スケールファクタ抽出手段 103 は、各サブバンド信号に対するスケールファクタを計算し、各サブバンドダイナミックレンジをそろえる。グループ分け手段 111 は、上記分割された 32 個の周波数成分を、外部からの制御情報である処理量制御情報 121 で指定されたグループ数に分割する。本実施の形態 1 では、図 5 に示すように、グループ数を 3 とし、各グループを周波数軸方向に連続したサブバンド信号として、0～15 サブバンドの低帯域グループ A、16～29 サブバンドの高帯域グループ B、及びビット割り当てを行わない 30～

31 サブバンドの無効グループ C にグループ分けを行う。なお、上記処理量制御情報 121 には C P U 負荷監視情報 21 と符号化モード設定信号 26 の情報が含まれているものとする。また、本実施の形態 1 では、上記各サブバンドグループにビットを割り当てるビット割り当て手段として、人間の耳に対して感度のよい低帯域に、M P E G によって指定された聴覚心理モデルに基づく信号対マスク比値との関係を使用して、高効率にビット割り当てを行う聴覚心理モデルビット割り当て手段 110 を使用し、人間の耳に対して比較的感度の低い高帯域に、スケールファクタ抽出手段 103 からのスケールファクタ情報に予め設定された各帯域毎の最小可聴限界値を加えたものを使用して、聴覚心理モデルビット割り当て方法と比較して低負荷でビット割り当てを行う帯域出力適応ビット割り当て手段 109 を使用するように構成している。

【0046】また、ビット割り当て処理手段 112 は、入力されたデジタルオーディオ信号を、聴覚心理モデルビット割り当てを行う 0～15 サブバンドの低帯域グループ A に対して、必要とされる聴覚心理分析を行うために、F F T (高速フーリエ変換) 手段 104 にて高速フーリエ変換を施すように F F T 手段 104 を制御する。そしてこの変換結果を用いて、聴覚心理分析手段 105 は、人間の聴覚の特性を利用した聴覚心理モデルに基づく信号対マスク比 (S M R) 値の関係モデルを導き出す。

【0047】また、割り当て可能ビット演算手段 113 は、サンプリング周波数や符号化のビットレート値から、確定するグループ全体に対する割り当て可能ビット数を、ビット割り当ての対象となる各グループのグループ全体に対する割合に、各グループに属する各帯域毎の特性に基づいた重み付けを加えたものを用いて、各グループ毎に独立したビット割り当て手段に対する割り当て可能ビット数を演算する。本実施の形態 1 では、聴覚心理モデルビット割り当て手段 110、及び帯域出力適応ビット割り当て手段 113 に対し、スケール・ファクタ・インデックス値、及び低帯域／高帯域の領域の比を考慮し、全体の割り当て可能ビット数から、両手段 110、113 への各割り当て可能ビット数を決定する。すなわち、実際には、スケールファクタ抽出手段 103 により求められた各スケール・ファクタ・インデックス値 $scf_index [i]$ から、下記の式 1、式 2 に示すように、両領域内の $scf_index [i]$ の加算値 V_{psy} 、 V_{non} を算出する。

【数 1】

$$V_{psy} = \sum_{i=0}^{psw_end-1} scf_index[i]$$

【数 2】

$$V_{non} = \sum_{i=psy_end}^{subband_end} scf_idx[i]$$

ここで、

`psy_end=16`: 聴覚心理モデルビット割り当てを行うサブバンド数
`subband_end=30`: 全ビット割り当てサブバンド数である。

【0048】次いで、人間の耳に感度のよい低帯域に、よりビットを多く配分するために、`Vpsy`に対し重み付けて行い、

$$V_{psy} = V_{psy} * 0.75$$

聴覚心理モデルビット割り当て可能数`psy_num`、及び帯域出力適応ビット割り当て可能数`non_num`を下記の式に基づいて求める。

$$V_{non} = V_{non} * psy_ratio$$

$$psy_num = all_alloc_num * V_{non} / (V_{psy} + V_{non})$$

$$non_num = all_alloc_num - psy_num$$

ここで、

`all_alloc_num`: 全体の割り当て可能ビット数
`psy_ratio`: `psy_end / (subband_end - psy_end)`である。

【0049】グループ毎の割り当て可能ビット数の範囲内(`psy_num, non_num`)で、聴覚心理モデルビット割り当て手段110は、聴覚心理分析手段105からのSMR値の関係モデルを利用し、低帯域グループAである0～15サブバンドに対してビット割り合てを行う。一方、帯域出力適応ビット割り当て手段109においては、高帯域グループBである16～29サブバンドに対してビット割り合てを行う。また、無効グループCである30～31サブバンドに対しては、無効なサブバンドとしたため、ビット割り当ては行われない。

【0050】これらのビット割り当て手段により決定された各サブバンド信号へのビット割り当て数に応じて、各サブバンド信号を量子化／符号化手段106において量子化／符号化し、補助情報符号化手段107により符号化されたヘッダ情報と補助情報をともに、ビットストリーム形成手段108においてビットストリームを形成して出力する。

【0051】処理量制御情報121からの情報が、例えば、符号化処理量を減少させるための情報であった場合には、図6に示すように、処理量の多い聴覚心理モデルビット割り当て手段110のビット割り当ての対象となる0～15サブバンド低帯域グループAのバンド幅を、0～7サブバンド低帯域グループA'に減少させ、反対に処理量の小さい帯域出力適応ビット割り当て手段109のビット割り当ての対象となるバンド幅を、8～29サブバンド低帯域グループB'となるように増加させる。さらに、符号化処理量の減少を考えた時の最終的な形態は、帯域出力適応ビット割り当て手段109のビッ

ト割り当てる対象を0～29サブバンドグループとしていることで処理量を制御する。この場合においては、実質的に聴覚心理モデルビット割り当て手段110は動作しないため、FFT手段104、及び聴覚心理分析手段105も動作しないことになる。

【0052】一方、処理量制御情報121からの情報が、例えば、符号化データの高音質化を図るために情報であった場合には、高効率(高音質)なビット配分が可能な聴覚心理モデルビット割り当て手段110のビット割り当てる対象となるバンド幅を増加させる。さらに、高音質を考えた時の最終的な形態は、聴覚心理モデルビット割り当て手段110のビット割り当てる対象を0～29サブバンドグループとしている。本実施の形態では以上のようにサブバンドグループの増減やビット割り当てる手段の切り替えを、オーディオ信号に復号可能な最小単位であるフレーム単位で行うことで、リアルタイムに符号化処理量を制御可能としている。

【0053】次に本実施の形態1による音声符号化装置の全体的な動作の流れについて図7を参照しつつ説明する。まず、図7(a)に示されるような構成を用い、各エンコーダ(23, 24)の処理負荷を認識するために、各エンコーダの各モード(ビット割り当てる対象となるバンド幅の変化)において、所定時間分のダミーデータを符号化することにより、CPU負荷監視部700は各モードでのCPU負荷値をデータテーブル701に格納する。そして、サンプル(データ)が入力されると、図7(b)のステップS70において、サブバンド分析が行われて32の周波数成分に分割され、続いてステップS71において各サブバンド信号のスケールファクタが計算される。

【0054】次いで、ステップS72において、CPU負荷の検出データがあるかどうかの判定が行われ、ここでは、動作開始直後なのでCPU負荷検出データはなく、従ってステップS74に進んで最も高音質な音声再生を行うことができる通常のグループ分けを行い、ステップS75に進んで聴覚心理モデルビット割り当てる処理が行われる。そして、ステップS76に進んで量子化／符号化処理が行われ、さらにステップS79においてビットストリーム形成が行われて一連の処理が終了するとともに、処理終了時に、入力された所定数のサンプルの符号化に要した時間をCPU負荷監視部700に通知し、現在のCPUの負荷が検出される。

【0055】すると、次回からの処理に際しては、ステップS72において、CPU負荷検出が「有」と判定されるようになり、ステップS73において、検出されたCPU負荷が実時間での符号化が不可能であると判定された場合には、ステップS77に進んで、データテーブル701を参照して最適なモード(グループ分け)を選択し、ステップS78における帯域ビット割り当てる処理を併用してステップS75における聴覚心理モデルビット割り当てる処理が実行される。

ト割り当て処理をそれぞれ、所定の割合にて行い、ステップS76に進んで量子化／符号化処理が行われ、ステップS79において、これら符号化されたデータを用いたビットストリームが形成されることになる。

【0056】なお、本実施の形態1における割り当て可能なビット演算手段113の演算を、スケール・ファクタ・インデックス値、及び低帯域、高帯域の領域の比を考慮し、各グループ毎に独立したビット割り当て手段に対する割り当て可能なビット数を演算するものとしたが、スケール・ファクタ・インデックス値の代わりにFFT手段104からの各グループに属するスペクトル信号レベルとしてもよく、各帯域毎に予め設定した最小可聴限界値としてもよい。

【0057】また、実施の形態1における符号化器101の処理量を制御するための情報を、符号化器101の内部に、CPUの処理量を監視するCPU負荷監視手段700を設け、CPUの処理能力を超えないように符号化器101を動作させるように構成したが、ユーザ入力などによる外部からの制御情報としてもよい。ユーザ入力をを行うことにより、音質、画質をユーザ好みに応じて優先させたエンコード処理を行うことが可能となる。

【0058】また、実施の形態1における符号化器101のビット割り当て手段として、人間の耳に対して感度のよい低帯域に、高効率にビット割り当てを行う聴覚心理モデルビット割り当て手段110を用い、高帯域に、低負荷でビット割り当てを行う帯域出力適応ビット割り当て手段109を固定的に用いるようにしたが、スケールファクタ抽出手段103からの信号により、各グループに属するサブバンド信号レベルが、予め設定した各帯域のしきい値以下の場合には、つまり、図8に示すように、低帯域に符号化データとして意味のある信号が高帯域に比べ少ない場合などには、ビット割り当て手段を帯域に応じて固定的に用いる必要はなく、高帯域に聴覚心理モデルビット割り当て手段110を用いるようにしてもよい。

【0059】また、図9に示すように、スケールファクタ抽出手段103からの信号により、各グループに属するサブバンド信号レベルと閾値とを比較判断するのではなく、スケールファクタ抽出手段103からの信号よりも分解能（周波数の）の高い、FFT手段104からの信号をビット割り当て処理制御手段112へ入力して、各グループに属するサブバンド信号のレベルと、予め設定した各帯域の閾値との比較判断を行うように構成してもよい。

【0060】（実施の形態2）次に本発明の実施の形態2によるデータ記録媒体について説明する。上記実施の形態1で示した音声符号化装置あるいは符号化方法の構成を実現するための符号化プログラムを、フロッピーディスク等のデータ記憶媒体に記録するようにすることにより、本発明の各実施の形態で示した処理を、独立した

コンピュータシステムにおいて簡単に実施することが可能となる。

【0061】すなわち、図10は、上記実施の形態1の符号化処理を、上記符号化プログラムを格納したフロッピーディスクを用いて、コンピュータシステムにより実施する場合を説明するための図である。図10(a)は、フロッピーディスクの正面からみた外観、断面構造、及びフロッピーディスク本体を示し、図10(b)は、該フロッピーディスク本体の物理フォーマットの例を示している。

【0062】上記フロッピーディスクFDは、上記フロッピーディスク本体DをフロッピーディスクケースFC内に収容した構造となっており、該フロッピーディスク本体Dの表面には、同心円状に外周からは内周に向かって複数のトラックTrが形成され、各トラックTrは角度方向に16のセクタSeに分割されている。従って、上記プログラムを格納したフロッピーディスクFDでは、上記フロッピーディスク本体Dは、その上に割り当てられた領域（セクタ）Seに、上記プログラムとしてのデータが記録されたものとなっている。

【0063】また、図10(c)は、フロッピーディスクFDに対する上記プログラムの記録、及びフロッピーディスクFDに格納したプログラムを用いた音声符号化処理を行うための構成を示している。上記プログラムをフロッピーディスクFDに記録する場合は、コンピュータシステムCsから上記プログラムとしてのデータを、フロッピーディスクドライブFDDを介してフロッピーディスクFDに書き込む。また、フロッピーディスクFDに記録されたプログラムにより、上記音声符号化装置をコンピュータシステムCs中に構築する場合は、フロッピーディスクドライブFDDによりプログラムをフロッピーディスクFDから読み出し、コンピュータシステムCsにロードする。

【0064】なお、上記説明では、データ記録媒体としてフロッピーディスクを用いて説明を行ったが、光ディスクを用いても上記フロッピーディスクの場合と同様にソフトウェアによる音声符号化処理を行うことができる。また、記録媒体は上記光ディスクやフロッピーディスクに限るものではなく、ICカード、ROMカセット等、プログラムを記録できるものであればよく、これらの記録媒体を用いる場合でも、上記フロッピーディスク等を用いる場合と同様にソフトウェアによる音声符号化処理を実施することができる。

【0065】（実施の形態3）次に本発明の実施の形態3による音声符号化方式、音声符号化装置について、図面を参照しながら説明する。図2で示した高帯域符号化処理手段24の構成としては、例えば、図3に示したような構成を用いる。また、低帯域符号化処理手段23の構成としては、例えば、図12示すように、図11で示したのと同様に、各帯域電力の偏在を利用して各帯域

(サブバンド) 每に符号化を行うような符号化方式を用いるが、所定の聴覚心理モデル分析手段のみを利用した各サブバンド信号に対するビット配分は行わず、新たに処理量の少ない簡易聴覚心理モデル部 4062 を設け、前フレームに出力された聴覚心理モデル部 4061 のマスキング閾値と該当フレームの帯域分割信号に基づき生成されたビット割り当て情報からビット配分を行うことを可能としている。

【0066】すなわち、図12は図2に示した低帯域符号化処理手段23の詳細な構成を示すプロック図であり、401は符号化器であり、後述するサブバンド分析手段402、スケールファクタ抽出手段403、ビット削り当て処理制御手段404、FFT処理手段405、聴覚心理分析手段406、聴覚心理モデルビット割り当て手段407、量子化／符号化手段408、補助情報符号化手段409、ビットストリーム形成手段410から構成されている。

【0067】以下、動作について説明する。上記サブバンド分析手段402は、入力されたデジタルオーディオ信号を32個の周波数成分に分割する。スケールファクタ抽出手段403は、各サブバンド信号に対するスケールファクタを計算し、各サブバンドダイナミックレンジをそろえる。FFT処理手段405は、入力されたデジタルオーディオ信号に対して、高速フーリエ変換を施す。聴覚心理分析手段406は、例えば、MPEGによって指定された通常聴覚心理モデル部4061、及び上記通常聴覚心理モデル部4061に比べて処理量の少ない上記簡易聴覚心理モデル部4062からなり、各モデルは信号対マスク比を算出する。

【0068】なお、上記通常聴覚心理モデル部4061は、下記の式3に基づき各サブバンド信号の信号対マスク比を算出するのに対し、上記簡易聴覚心理モデル部4062は、下記の式4に示すように、該当フレームにおいては、各サブバンドの最小マスキング・レベルの算出は行なわず、上記通常聴覚心理モデル部4061によって算出された最近の前フレームでの最小マスキング・レベルを用い、音圧は該当フレームのスケールファクタ抽出手段403によって抽出されたスケール・ファクタ値を用いて信号対マスク比を算出する。

【数3】

$$SMR_{sb}(n) = L_{sb}(n) - LT_{min}(n) \quad db$$

ここで、

$L_{sb}(n)$: 各サブバンドの音圧

$LT_{min}(n)$: 各サブバンドの最小マスキング・レベルである。

【数4】

$$SMR_{sb}(n) = L_{sb}(n) - LT_{min}(n) \quad db$$

ここで、

$$L_{sb}(n) = 20 \cdot \log(scf_{max}(n) + 32768) - 10 \quad db$$

$scf_{max}(n)$: 該当フレームの各サブバンドに対するスケール・ファクタ値

$LT_{min}(n)$: 最近の上記通常聴覚心理モデル部4061にて算出された各サブバンドの最小マスキング・レベル

【0069】ビット割り当て処理制御手段404は、処理量制御情報121の情報を基に、本実施の形態3では、図13に示すように、Nを3とし、低負荷処理が実現可能な上記簡易聴覚心理モデル部4062と高音質化を実現することができる最適なビット配分情報を出力することが可能な上記通常聴覚心理モデル部4061を、何フレームに一度行なうかの制御、及びFFT処理手段405における高速フーリエ変換を行なうか否かの制御を行なう。例えば、図13の状態において、処理量制御情報121の情報を基に、符号化処理に割くCPU占有率を下げるという情報が、上記ビット割り当て処理制御手段404に知らされると、処理量の小さい上記簡易聴覚心理モデル部4062の使用を多くするために、Nの値を大きくする。反対に、符号化処理に割くCPU占有率をもっと使用してもよいという情報が、上記ビット割り当て処理制御手段404に知らされると、高音質化が実現可能な上記通常聴覚心理モデル部4061の使用を多くするためにNの値を小さくする。これにより、処理量の制御を可能とすることができる。

【0070】聴覚心理モデルビット割り当て手段407は、上記ビット割り当て処理制御手段404からの情報である信号対マスク比の関係から上記サブバンド分析手段402により分割された各サブバンド信号に対し、ビットの割り当てを行なう。量子化／符号化手段408により、各サブバンド信号の量子化、及び符号化を行い、補助情報符号化手段409からの補助データと共にビットストリーム形成手段410によりビットストリームが形成され出力される。

【0071】このように本実施の形態3によれば、ビット割り当てをNフレームに一度の割合で行なうようにしたので、時間軸方向でのCPU負荷を低減することができるようになる。

【0072】なお、ここでは仮に符号化器401を、図2に示した低帯域符号化処理手段23としたが、低帯域信号に対してのみ適用されるのではなく、全帯域信号に対して適用するようにしてもよい。

【0073】(実施の形態4) 次に本発明の実施の形態4による音声符号化方式、音声符号化装置について、図面を参照しながら説明する。図14は、図11で示したのと同様に、各帯域電力の偏在を利用して各帯域(サブバンド)毎に符号化を行なうような符号化方式を用いているが、出力ビットストリーム中に、オーディオデータ以外の外部データを付加する機能を備えている点が異なる。上記外部データとしては、画像データやテキストデータなどが想定される。

【0074】すなわち、図14に示す符号化器501

は、後述するサブバンド分析手段 502、スケールファクタ抽出手段 503、FFT処理手段 504、聴覚心理分析手段 505、ビット割り当て手段 506、量子化／符号化手段 507、補助情報符号化手段 508、ビットストリーム形成手段 509、ビット割り当て処理制御手段 510、付加データ符号化手段 511 から構成されている。

【0075】以下、動作について説明する。上記サブバンド分析手段 502 は、入力されたデジタルオーディオ信号を 32 個の周波数成分に分割する。スケールファクタ抽出手段 503 は、各サブバンド信号に対するスケールファクタを計算し、各サブバンドダイナミックレンジをそろえる。FFT 処理手段 504 は、入力されたデジタルオーディオ信号に対して、高速フーリエ変換を施す。聴覚心理分析手段 505 は、例えば、MPEG によって指定された聴覚心理モデルにより、信号対マスク比を算出する。

【0076】ビット割り当て処理制御手段 510 は、出力カビットストリーム中に付加するデータを一時的に格納しておく付加データバッファ 512 を監視し、付加データがあるか否かの判断、あるいは付加データがオーバフローするか否かの判断から生成された割り当て範囲制御情報 513 を基に、ビット割り当て手段 506 に対し、ビット割り当てを行なう範囲を指定する。

【0077】例えば、付加データバッファ 512 にデータが存在しない場合には、図 15 に示すように、サブバンド 0～29 に対してビット割り当てが行われる。この場合、全体の割り当て可能ビット数を 100 とし、サブバンド 0～15 に 80 ビット、サブバンド 16～29 に 20 ビットが割り当てられている。

【0078】そして、付加データバッファ 512 に外部よりデータが書き込まれ、付加データが存在する状態になった場合には、つまり、割り当て範囲制御情報 513 として、付加データを挿入するという指示が、ビット割り当て処理制御手段 510 に対して知らされ、本実施の形態 4 では、例えば、サブバンド 0～15 に 80 ビットを割り当て、本来割り当てられるべきサブバンド 16～29 に対してはビット割り当てを行なわず、余った 20 ビットをデータの付加ビット数として割り当てられる。また、ビット割り当てが行われないサブバンド 16 以降のサブバンドに対しては、処理量を削減するために、該当範囲の FFT 处理、及び聴覚心理分析を行なわないようにしてよい。

【0079】そして、上記ビット割り当てが行われたサブバンドに対し、量子化／符号化手段 507 により、量子化、及び符号化を行い、補助情報符号化手段 508 からの補助データと、例えば、MPEG のアンシラリーデータとして符号化された付加データと共に、ビットストリーム形成手段 509 によりビットストリームが形成され出力される。

【0080】このように本実施の形態 4 によれば、定ビットレートでの伝送を行う際に、オーディオデータ以外の付加データの量に応じて、符号化時のビット割り当て範囲を制御して符号化するオーディオデータの量を可変として符号化データストリーム中に付加データを挿入するようにしたので、余剰帯域に様々なデータを重畠して帯域を有効に利用することができる。なお、ビット割り当て処理制御手段 510 により実行されるビット割りあて範囲の制御は、フレーム単位で行い、そのビット割り当て範囲も付加データバッファ 512 のデータ量に応じて、可変可能としている。

【0081】これらの処理により、付加データ挿入においてもビット割り当て範囲内の音質を損なうことなく、リアルタイムにデータ挿入量の制御を可能とすることができる。

【0082】(実施の形態 5) 次に本発明の実施の形態 5 による音声符号化方式、音声符号化装置について、図面を参照しながら説明する。図 16 は本実施の形態 5 による音声符号化方式を用いた音声符号化装置の符号化器の構成を示すブロック図であり、図において、図 2 と同一符号は同一または相当部分を示し、160～162 はそれぞれ独立的に動作可能な符号化処理手段 A～C、163 は各符号化処理手段 A～C の処理負荷値情報を格納するための処理負荷値格納バッファ、164 は上記各符号化処理手段 A～C にサンプルとなるデータを供給するためのサンプルデータバッファである。

【0083】次に動作について説明する。符号化処理を行う前の初期化時に、まず、サンプルデータバッファ 164 に格納されている所定のサンプルデータを各符号化処理手段 A～C に供給し、これによって発生する符号化処理手段 A～C、あるいは聴覚心理モデルの処理負荷値を処理負荷値格納バッファ 163 に格納する。

【0084】そして、上記処理負荷値を昇順、もしくは降順にて出力することによって、装置にて使用される CPU の性能に見合った符号化処理手段を迅速に選択し、当該符号化処理手段によって符号化処理を行う。符号化処理の内容については、実施の形態 1 で示したものと同じであるので個々では省略する。

【0085】このように本実施の形態によれば、符号化処理前の初期化時に、サンプルデータを用いて各符号化処理手段を動作させ、そのときの負荷値を取得して、使用する CPU の処理能力に適した符号化処理手段を選択して用いるようにしたので、CPU の負荷が減少して、最適な符号化処理を行うことができるようになる。

【0086】なお、以上の各実施の形態では、音声符号化装置として、PC を用いて実現する構成を例に挙げて説明したが、例えば VTR カメラや DVD エンコーダなどの機器に組み込んで用いるような場合にも適用することができる。

【0087】また、上記各実施の形態では、音声のみを

取り扱うようにしたが、音声とともに映像を処理する場合には、図17に示すように、図2の構成において、音声信号とは別に映像信号を入力し、低帯域符号化処理手段と高帯域符号化処理手段に代えて、映像符号化処理手段170と音声符号化処理手段171を設け、さらに、ビットストリーム形成処理手段に代えてシステムストリーム処理手段172を設けた構成とすることにより、対応することができる。以上のような構成を用いて、外部からの制御情報に基づいて、上記各実施の形態で説明したような方法で、音声符号化の演算量を変更したり、演算量の異なる複数の音声符号化方式を切り替えたりすることによって、CPUとしての全体の演算量を制御することが可能となる。また、あるいは、符号化すべき音声信号の量に応じて、符号化する映像信号の処理量を変化させるように構成してもよい。

【0088】さらに、サブバンド方式のコーディングを行うMPEG1以外に、時間／周波数変換を行うMPEG2、AAC、Dolby AC-3、ATRAC(MD)などのコーディング方式を行う場合についても、符号化処理に関わる各手段を、図18に示すように、演算量の異なる第1の量子情報算出手段181と第2の量子化情報算出手段182に置換するとともに、これらを量子化手段制御手段180によって選択して使用し、符号化情報の代わりに量子化情報を取り扱う構成とすることで、同様に対応することが可能である。

【0089】

【発明の効果】以上のように、この発明の請求項1にかかる音声符号化方式によれば、デジタルオーディオ信号を複数の周波数帯域に分割し、各帯域毎に符号化を行う音声符号化方式であって、上記分割された各帯域に対するビット割り当て情報を生成し、それぞれ処理量の異なるビット割り当て手段を複数有し、外部からの制御情報に基づいて、上記複数のビット割り当て手段の中から、所定のビット割り当て手段を用いて処理がなされるように、使用するビット割り当て手段を切り替えてビット割り当てを実行して、符号化を行うようにしたので、常に最適な処理量のビット割り当て手段を選択して使用することができ、稼動状態において占有できるCPUの処理量を超えないような符号化が可能となり、リアルタイムエンコード時に処理が入力信号に対して間に合わないということがない、つまり再生音に音切れがない符号化を行うことができるという効果がある。

【0090】また、本発明の請求項2にかかる音声符号化方法によれば、上記外部からの制御情報として、符号化処理を行うために占有できる中央演算処理装置の処理量を表す負荷値を用い、上記負荷値に基づいて、上記中央演算処理装置上で符号化処理において各ビット割り当て手段を用いて符号化動作を行った時の各処理量を予め記憶したデータテーブルを参照して、上記符号化処理が占有できる中央演算処理装置の処理量を超えないようビ

ット割り当て手段の選択を行うようにしたので、中央演算処理装置は常に稼動能力を超えるような要求を受けることが無くなり、システム全体の制御をスムーズに行うことができるという効果がある。

【0091】また、本発明の請求項3にかかる音声符号化方式によれば、上記請求項2記載の行う音声符号化方式において、上記負荷値として、符号化処理を行うために占有可能な上記中央演算処理装置の処理量を監視する監視手段からの処理量制御情報を用いるとしたので、占有可能な中央演算処理装置の最高パフォーマンスの範囲内で、最適な処理量のビット割り当て手段を選択することができ、リアルタイムエンコード時に処理が入力信号に対して間に合わないということがない、つまり再生音に音切れがない符号化を行うことができるという効果がある。

【0092】また、本発明の請求項4にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、上記ビット割り当て手段によるビット割り当て処理として、符号化データの高音質化を実現可能な高効率にビット割り当てを行う高効率ビット割り当て方法を用いた処理と、該高効率ビット割り当て方法を用いた処理と比較して処理量の少ない低負荷でビット割り当てを行う低負荷ビット割り当て方法を用いた処理を行うようにしたので、符号化器が符号化データの高音質化を優先するか、又は音質よりも符号化処理の低負荷性を優先するかの処理を適宜切り替えて実行することができる符号化を実現することができるという効果がある。

【0093】また、本発明の請求項5にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、上記符号化時に使用されるビット割り当て手段の切り替えを、オーディオ信号に復号可能な最小単位であるフレーム単位で行うようにしたので、リアルタイムエンコード時に、動作CPU上で該CPUを共有する他のアプリケーションなどのCPU占有率が突然大きくなつた場合などにおいても、フレーム単位時間で符号化処理が占有できるCPUの処理量に追従可能となり、また、リアルタイムで音質や処理量を制御可能とすることができるという効果がある。

【0094】また、本発明の請求項6にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、複数の周波数帯域に分割された各帯域のサブバンド信号を、各々予め定められた所定個数のサブバンド信号からなるグループとなるようにグループ分けを行い、各グループに対して独立したビット割り当て処理を行い、各帯域に対するビット割り当て情報を生成するようにしたので、各帯域毎の特性に応じたビット割り当て処理を選択して符号化を行うことができるという効果がある。

【0095】また、本発明の請求項7にかかる音声符号化方式によれば、上記請求項6記載の音声符号化方式に

において、上記グループ分けが、グループの数、又はグループ内の周波数軸方向に連続したサブバンド信号の数を、上記外部からの制御情報により指定された数、又は上記監視手段からの処理量制御情報に基づいて指定された数となるように可変的に行われるようとしたので、CPUの使用状況に応じてダイナミックにグループ分けを行うことができるという効果がある。

【0096】また、本発明の請求項8にかかる音声符号化方式によれば、上記請求項7記載の音声符号化方式において、上記サブバンド信号の数の変更処理を、オーディオ信号に復号可能な最小単位であるフレーム単位で行うようにしたので、ビット割り当て方式の変更をきめ細かく行うことができ、高精度な符号化器を実現することができるという効果がある。

【0097】また、本発明の請求項9にかかる音声符号化方式によれば、請求項8記載の音声符号化方式において、上記グループ分け時に、ビット割り当てを行わないグループを少なくとも1つ設けるようにしたので、オーディオ信号に復号可能な最小単位であるフレーム単位で、グループの数、またはグループ内の周波数軸方向に連続したサブバンド信号の数を、外部からの制御情報により指定された数、または監視手段からの処理量制御情報により指定された数に変えることで、ビット割り当てが行われないグループに属する帯域の信号を符号化処理する必要がなくなり、また、ビット割り当てが行われないグループに属する帯域に割り当てられるべきビットを、ビット割り当てが行われる他のグループの帯域に分配することができ、その結果、符号化処理が占有するCPUの処理量を制御可能となるとともに、ビット割り当てが行われる他のグループの帯域の音質を向上することができるという効果がある。

【0098】また、本発明の請求項10にかかる音声符号化方式によれば、上記請求項6記載の音声符号化方式において、上記サブバンド信号のグループ分けにより、低帯域に属するグループにグループ分けされたサブバンド信号に対し、符号化データの高音質化を実現可能な高効率ビット割り当て方法を用いて、グループ内のサブバンド信号にビット割り当てを行な一方、高帯域に属するグループにグループ分けされたサブバンド信号に対し、上記高効率ビット割り当て方法と比較して処理量の少ない低負荷ビット割り当て処理を用いて、グループ内のサブバンド信号にビット割り当てを行なうようにしたので、人間の耳に対して感度のよい低帯域については符号化データの高音質化を図ることができ、一方、人の耳に対して感度の悪い高帯域については処理量優先の低負荷ビット割り当てを用いられるようになり、全体として処理量を削減した符号化を行うことができるという効果がある。

【0099】また、本発明の請求項11にかかる音声符号化方式によれば、上記請求項6記載の音声符号化方

において、各グループ毎に独立したビット割り当て手段に対する割り当て可能ビット数を決定する割り当て可能ビット演算手段を設け、各グループのグループ全体に対する割合に各グループに属する各帯域毎の特性に基づいた重み付けを加味したものを用いて、グループ全体に対する割り当て可能ビット数を、各グループ毎に独立したビット割り当て手段に対し振り分けるようにしたので、入力信号、又は各帯域の特性に対して、聴覚特性を考慮した符号化データの高音質を実現するのに最適な各グループのビット割り当て手段に対するビット配分が可能な符号化を行うことができるという効果がある。

【0100】また、本発明の請求項12にかかる音声符号化方式によれば、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、各帯域毎の所定の最小可聴限界値に基づいた重み付けとすることにより、上記人間が聴く際に意味の有る効果的なビット割り当て処理を行うことができるという効果がある。

【0101】また、本発明の請求項13にかかる音声符号化方式によれば、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号にサブバンド分析を施して得られる、各グループに属する各周波数帯域のサブバンド信号レベルに基づいた重み付けとすることにより、効果的なビット割り当て処理を行うことができるという効果がある。

【0102】また、本発明の請求項14にかかる音声符号化方式は、上記請求項11記載の音声符号化方式において、各グループに属する各帯域毎の特性に基づいた重み付けを、入力デジタルオーディオ信号に線形変換を施して得られる、各グループに属するスペクトル信号レベルに基づいた重み付けとすることにより、効果的なビット割り当て処理を行うことができるという効果がある。

【0103】また、本発明の請求項15にかかる音声符号化方式によれば、上記請求項6記載の音声符号化方式において、各グループに属する信号レベルが、所定のしきい値以上の高レベルな信号に対しては、符号化データの高音質化を実現可能な高効率ビット割り当て方法を用いてビット割り当てを行い、各グループに属する信号レベルが、所定のしきい値以下の低レベルな信号に対しては、上記高効率ビット割り当て方法と比較して処理量の少ない低負荷ビット割り当て方法を用いてビット割り当てを行なうようにしたので、他の帯域に比べ符号化データとしてそれほど重要でない信号に対し、処理負荷を割くことなく、符号化データの高音質化を図ることができる符号化を行うことができるという効果がある。

【0104】また、本発明の請求項16にかかる音声符号化方式によれば、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルを、入力デジタルオーディオ信号にサブバンド分析を施して得

られるサブバンド信号レベルとすることにより、効果的なビット割り当て処理を行うことができるという効果がある。

【0105】また、本発明の請求項17にかかる音声符号化方式によれば、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルを、入力デジタルオーディオ信号に線形変換を施して得られるスペクトル信号レベルとしたので、することにより、効果的なビット割り当て処理を行うことができるという効果がある。

【0106】また、本発明の請求項18にかかる音声符号化方式によれば、上記請求項15記載の音声符号化方式において、上記各グループに属する信号レベルを、所定の各帯域毎の最小可聴限界値としたので、上記人間が聞く際に意味の有る効果的なビット割り当て処理を行うことができるという効果がある。

【0107】また、本発明の請求項19にかかる音声符号化方式によれば、上記請求項4, 10, 15のいずれかに記載の音声符号化方式において、上記符号化データの高音質化を実現可能な高効率ビット割り当て処理として、所定の聴覚心理モデルに基づく信号対マスク比値との関係を使用してビット割り当てを行ふものとし、一方、上記高効率ビット割り当て処理と比較して処理量の少ない低負荷ビット割り当て処理として、複数の周波数帯域に分割された信号レベルに各帯域毎の所定の最小可聴限界値を加味してビット割り当てを行ふものとしたので、全体として人間の耳で聞いた限りは音質を損なうことなく、システムの処理量を軽減することができるという効果がある。

【0108】また、本発明の請求項20にかかる音声符号化方式は、上記請求項19記載の音声符号化方式において、上記聴覚心理モデルがMPEGによって指定された聴覚心理モデルとすることにより、MPEG (Motion Picture Experts Group) を用いた音声符号化処理においても上記同様の効果を得ることができるという効果がある。

【0109】また、本発明の請求項21にかかる音声符号化方式は、上記請求項5または請求項8記載の音声符号化方式において、上記オーディオ信号に復号可能な最小単位であるフレームが、MPEG (Motion Picture Experts Group) によって指定されたフレームとすることにより、MPEGを用いた音声符号化処理においても上記同様の効果を得ることができるという効果がある。

【0110】また、本発明の請求項22にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、上記ビット割り当て手段は、分割化された各帯域に対し、所定の聴覚心理モデルから出力される情報に基づいてビット割り当て情報を生成するものであり、N(N=1, 2, 3...)フレームに1度、上記所定の聴覚心理モデルから出力される情報に基づいてビット割り当

て情報を生成し、上記ビット割り当て情報を生成しなかつたフレームに対しては、上記聴覚心理モデルから出力された情報と上記分割された各帯域の信号情報を基づいてビット割り当て情報を生成し、符号化を行なうようにしたので、時間軸方向でのCPU負荷を低減することができるという効果がある。

【0111】また、本発明の請求項23にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、段階的に処理量の制御可能な聴覚心理モデルを有し、外部からの制御情報に基づいて、上記聴覚心理モデルの処理量制御を行ない、所定の処理量の聴覚心理モデルを用いて処理がなされるよう、各帯域に対するビット割り当て情報を生成するようにしたので、聴覚的な効果を加味したCPU負荷制御を行うことができるという効果がある。

【0112】また、本発明の請求項24にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式において、それぞれ処理量の異なる聴覚心理モデルを複数有し、外部からの制御情報に基づいて、上記複数の聴覚心理モデルの中から、所定の聴覚心理モデルを用いて処理が成されるように、使用する聴覚心理モデルを切り替えて、各帯域に対するビット割り当て情報を生成するようにしたので、より簡単に聴覚的な効果を加味したCPU負荷制御を行うことができるという効果がある。

【0113】また、本発明の請求項25にかかる音声符号化方式によれば、デジタルオーディオ信号を複数の周波数帯域に分割し、該分割された各帯域に対するビット割り当て情報を生成し、所定ビットレートでの伝送を目的として各帯域毎に符号化を行なう音声符号化方式であって、符号化データストリーム中にデータを挿入するフレームにおけるビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御するようにしたから、CPU処理の負荷が一定となるとともに、余剰帯域に様々なデータを重畳して帯域を有效地に利用することができるという効果が得られる。

【0114】また、本発明の請求項26にかかる音声符号化方式によれば、上記請求項25記載の音声符号化方式において、外部からの制御情報に基づいて、フレーム単位でビット割り当て範囲を制御し、符号化オーディオデータ量を可変に制御するようにしたので、CPUの処理負荷を効率よく低減することができるという効果がある。

【0115】また、本発明の請求項27にかかる音声符号化方式によれば、上記請求項26記載の音声符号化方式において、上記外部からの制御情報として、付加データのバッファを監視する監視手段からのデータ量制御情報を用いるようにしたので、負荷データを優先的に重畳することができるという効果がある。

【0116】また、本発明の請求項28にかかる音声符号化方式によれば、上記請求項1記載の音声符号化方式

において、符号化処理が実行される中央演算処理装置の性能に応じて、符号化処理動作前の初期化時に複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報を外部へ出力するようにしたので、実際に符号化を行う前に、使用されるCPUの性能に関する情報を取得することができ、CPUの処理負荷を効率的に低減することができるという効果がある。

【0117】また、本発明の請求項29にかかる音声符号化方式によれば、上記請求項28記載の音声符号化方式において、外部への情報として出力される複数のビット割り当て手段、または複数の聴覚心理モデルの各処理負荷値情報が、降順、あるいは昇順で出力されるようにしたので、符号化処理手段の選択を迅速に行うことができるという効果がある。

【0118】また、本発明の請求項30にかかる音声符号化方式によれば、映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、異なる複数の演算量で符号化を行うものであって、上記音声信号もしくは映像信号の符号化の演算量を変更することで、上記中央演算処理装置上の処理の全体的な演算量を制御するようにしたので、音声とともに映像信号を符号化する処理においても、CPUの負荷処理を行うことができるという効果がある。

【0119】また、本発明の請求項31にかかる音声符号化方式によれば、映像信号、音声信号の符号化処理が同一の中央演算処理装置にて処理がなされる音声符号化方式において、演算量の異なる複数の符号化方式を用いて符号化を行うものであって、上記音声信号の符号化の符号化方式を変更することで、上記中央演算処理装置上の処理の全体的な演算量を制御するようにしたので、音声とともに映像信号を符号化する処理においても、CPUの負荷処理を行うことができるという効果がある。

【0120】また、本発明の請求項32にかかる音声符号化方式によれば、上記請求項30または31記載の音声符号化方式において、上記中央演算処理装置上の処理の制御を、外部からの制御情報に基づいて行うようにしたので、CPUの処理負荷を効率よく低減することができるという効果がある。

【0121】また、本発明の請求項33にかかる音声符号化方式によれば、デジタルオーディオ信号に対して、時間／周波数変換を施し、量子化情報を生成することで符号化を行う音声符号化方式であって、それぞれ演算量の異なる複数の量子化情報算出手段を有し、外部からの制御情報に基づいて、上記複数の量子化情報算出手段の中から、所定の量子化情報算出手段を選択して処理が行われるように、使用する量子化情報算出手段を切り替えて量子化情報を算出して符号化を行うようにしたので、時間／周波数変換方式のコーディングを行う符号化処理装置においても、CPUの処理負荷を低減することができるという効果がある。

【0122】また、本発明の請求項34にかかる音声符号化装置によれば、上記請求項1ないし請求項33のいずれかに記載の音声符号化方式を用いて音声符号化を行うようにしたので、該音声符号化方式を組み込んだVTRカメラなどの機器においても上記同様の効果を得ることができるという効果がある。

【0123】また、本発明の請求項35にかかるデータ記憶媒体は、上記請求項1ないし請求項33のいずれかに記載の音声符号化方式のステップが記録されているので、該記憶媒体を用いて上記音声符号化方式を装置に組み込むことにより、上記同様の効果を得ることができるという効果がある。

【図面の簡単な説明】

【図1】本発明の実施の形態1による音声符号化方式を用いた音声符号化装置を実現するために、パソコン用コンピュータを用いた全体的なシステムとしての構成を示すブロック図である。

【図2】上記実施の形態1による音声符号化装置を構成する符号化器の構成を示すブロック図である。

【図3】上記符号化器を構成する高帯域符号化処理手段の詳細な構成を示すブロック図である。

【図4】上記実施の形態1による音声符号化装置を構成する符号化器のより詳細な構成を示すブロック図である。

【図5】上記実施の形態1による音声符号化方式で使用する、各グループに対するビット割り当て処理の一例を示した模式図である。

【図6】上記実施の形態1による音声符号化方式で使用する、各グループに対するビット割り当て処理の他の一例を示した模式図である。

【図7】上記実施の形態1による音声符号化装置を構成する符号化器の符号化動作を説明するためのフローを示す図である。

【図8】上記実施の形態1による音声符号化方式で使用する、各グループに対するビット割り当て処理の、閾値を用いて処理を行う例を示した模式図である。

【図9】本発明の上記実施の形態1による音声符号化装置を構成する符号化器の変形例の詳細な構成を示すブロック図である。

【図10】本発明の実施の形態2によるデータ記憶媒体及び該記憶媒体を用いて音声符号化装置を構成する場合の構成を示すブロック図である。

【図11】従来の音声符号化装置を構成する符号化器の構成を示すブロック図である。

【図12】本発明の実施の形態3による音声符号化装置を構成する低帯域符号化処理手段の詳細な構成を示す図である。

【図13】上記実施の形態3による音声符号化装置による低帯域符号化時の各フレームにおける聴覚心理モデルの状態を説明するための図である。

【図14】本発明の実施の形態4による音声符号化装置を構成する低帯域符号化処理手段の詳細な構成を示す図である。

【図15】上記実施の形態4による音声符号化装置を用いたビット割り当て処理の一例を示す図である。

【図16】本発明の実施の形態5による音声符号化装置を構成する符号化器の構成を示すブロック図である。

【図17】音声信号とともに映像信号を取り扱う場合の符号化器の構成を示すブロック図である。

【図18】時間／周波数変換方式のコーディングを行う符号化処理装置における符号化処理において本発明を適用した場合の構成を示すブロック図である。

【符号の説明】

1 パーソナルコンピュータ（音声符号化装置）

1 1 HDD

1 2 a PDD

1 2 b FDD

1 3 メモリ

1 4 CPU（中央演算処理装置）

1 5 データバス

1 6 ビデオキャプチャーカード

1 7 カメラ

1 8 サウンドカード

1 9 マイク

2 0 符号化器

2 1 CPU負荷監視情報

2 2 符号化手段制御手段

2 3 低帯域符号化処理手段

2 4 高帯域符号化処理手段

2 5 ビットストリーム形成処理手段

2 6 符号化モード指定信号

1 0 1 符号化器

1 0 2 サブバンド分析手段

1 0 3 スケルファクタ抽出手段

1 0 4 FFT手段

1 0 5 聴覚心理分析手段

1 0 6 量子化／符号化手段

1 0 7 補助情報符号化手段

1 0 8 ビットストリーム形成手段

1 0 9 帯域出力適応ビット割り当て手段

1 1 0 聴覚心理モデルビット割り当て手段

1 1 1 グループ分け手段

1 1 2 ビット割り当て処理制御手段

1 2 1 処理量制御情報

1 6 0 ~ 1 6 2 符号化処理手段 A ~ C

1 6 3 処理負荷値格納バッファ

1 6 4 サンプルデータバッファ

1 7 0 映像符号化処理手段

1 7 1 音声符号化処理手段

1 7 2 システムストリーム形成処理手段

1 8 0 量子化手段制御手段

1 8 1 第1の量子化情報算出手段

1 8 2 第2の量子化情報算出手段

F C フロッピーディスクケース

F D フロッピーディスク

D フロッピーディスク本体

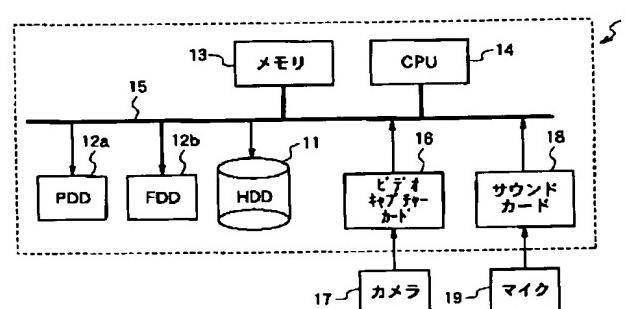
S e セクタ

T r トランク

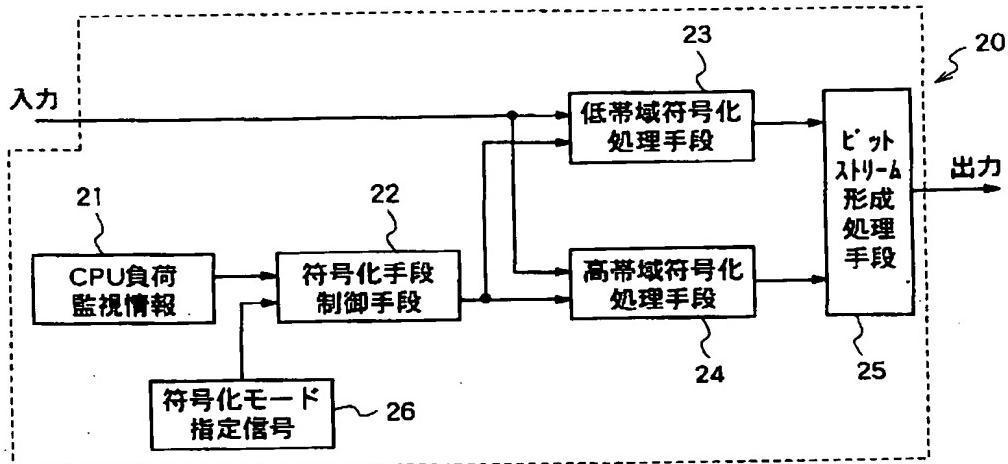
C s コンピュータシステム

F DD フロッピーディスクドライブ

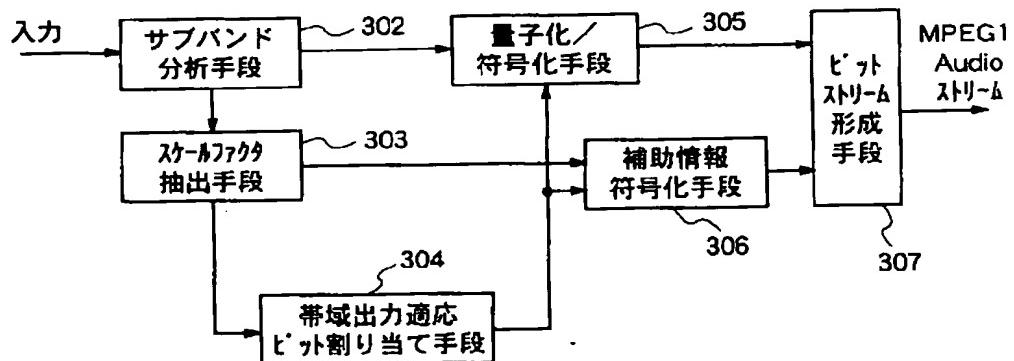
【図1】



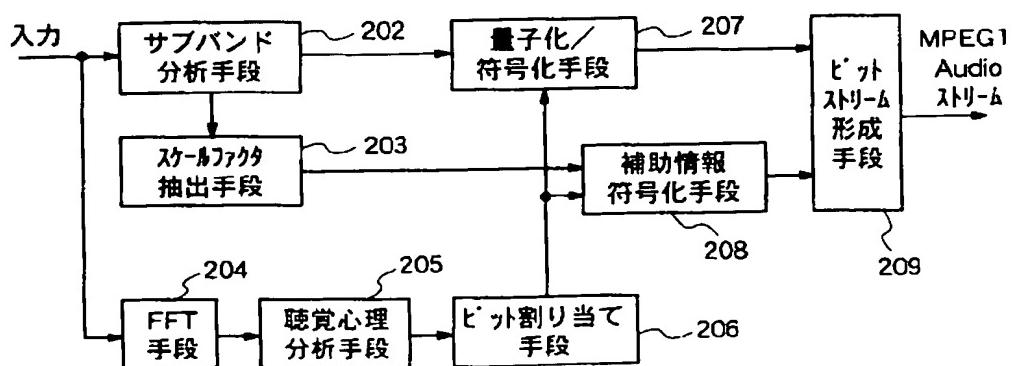
【図 2】



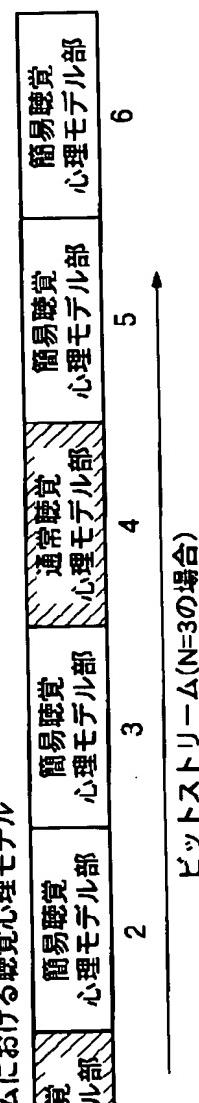
【図 3】



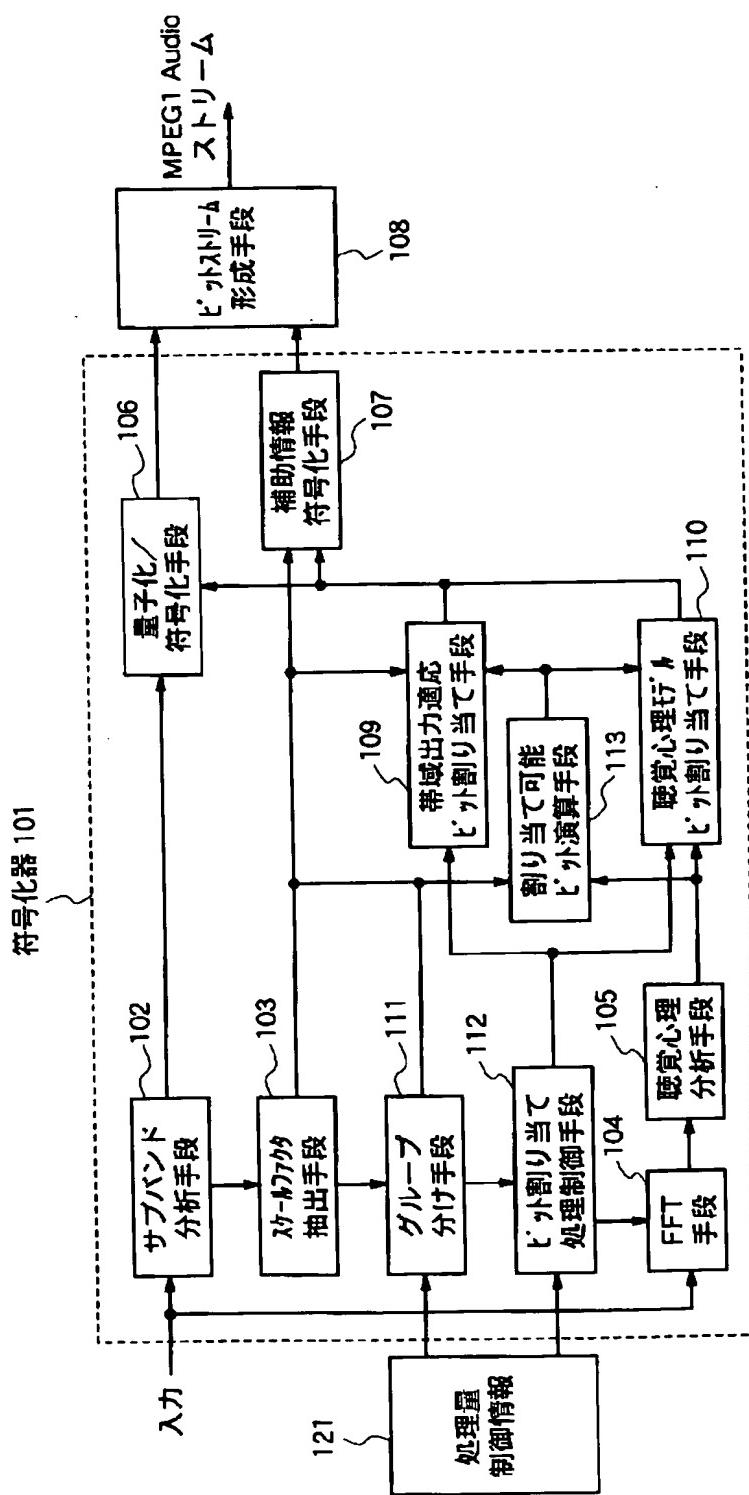
【図 1.1】



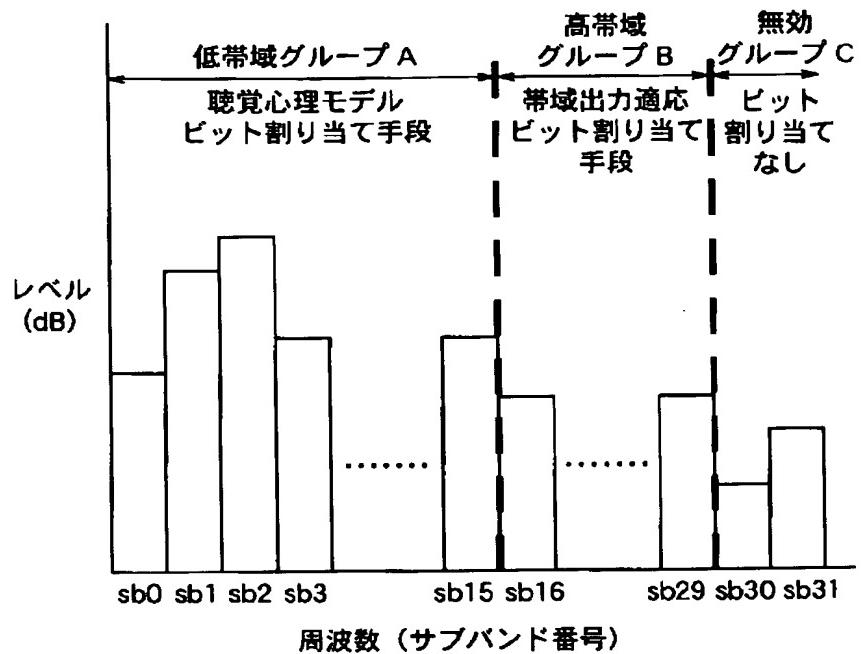
【図13】



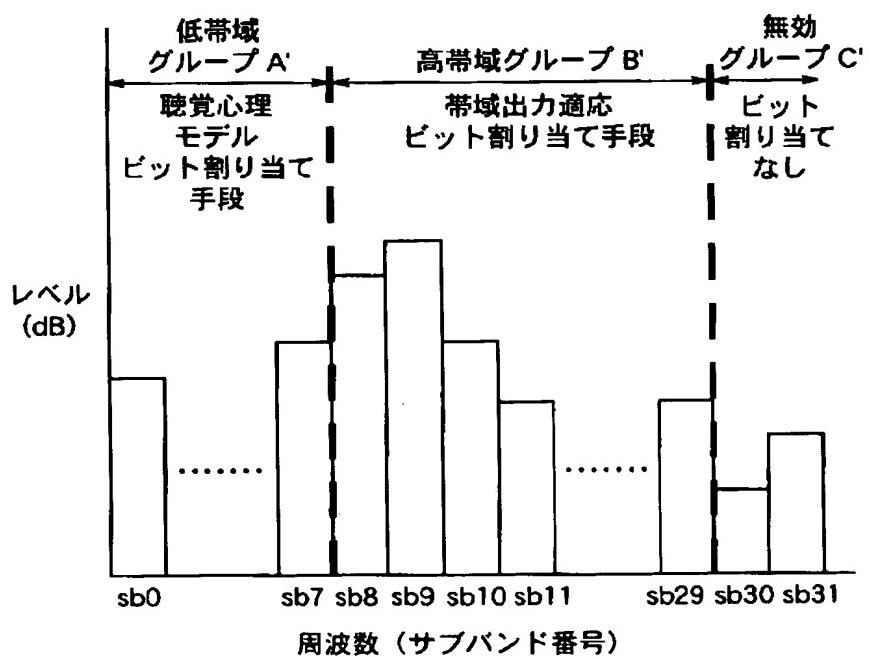
【図4】



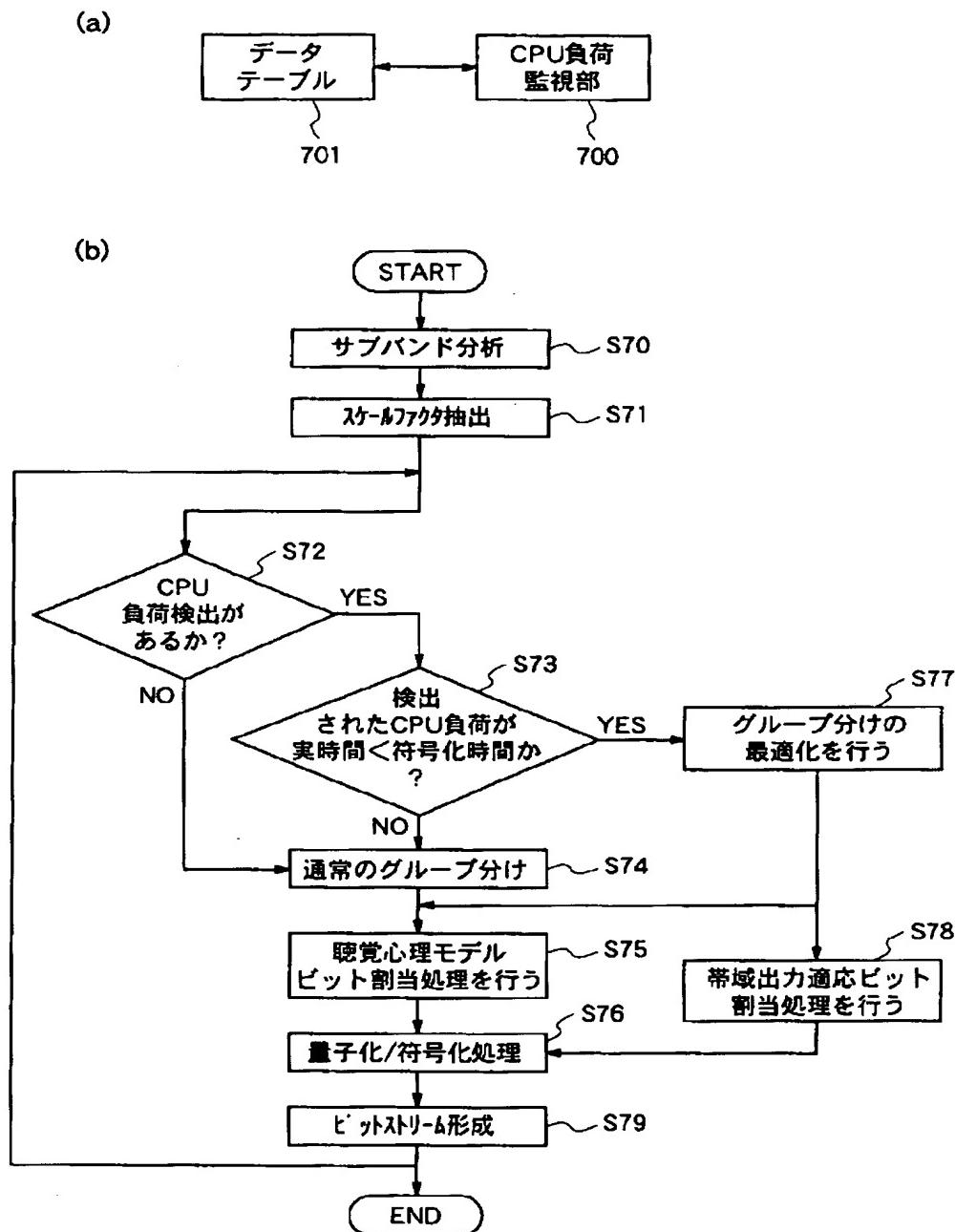
【図 5】



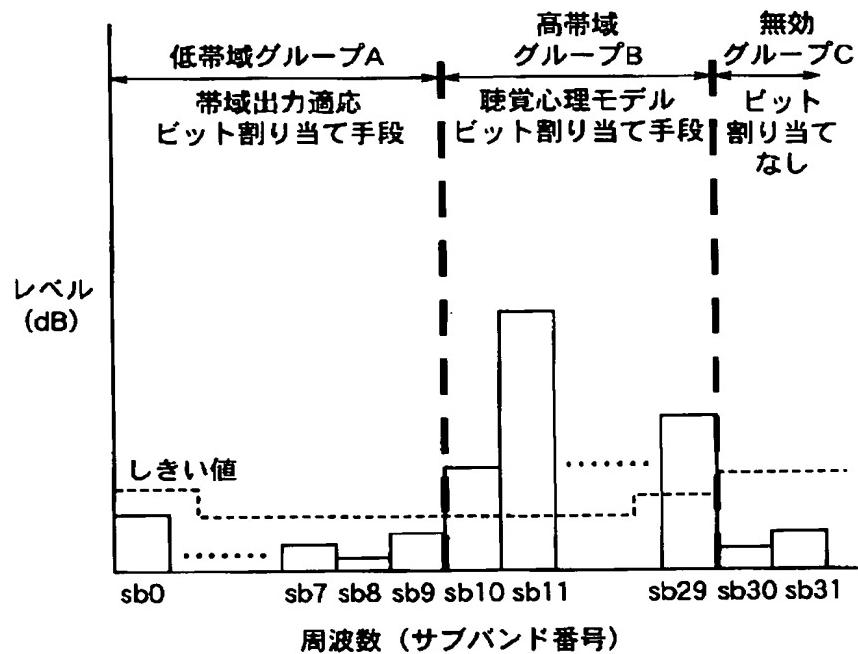
【図 6】



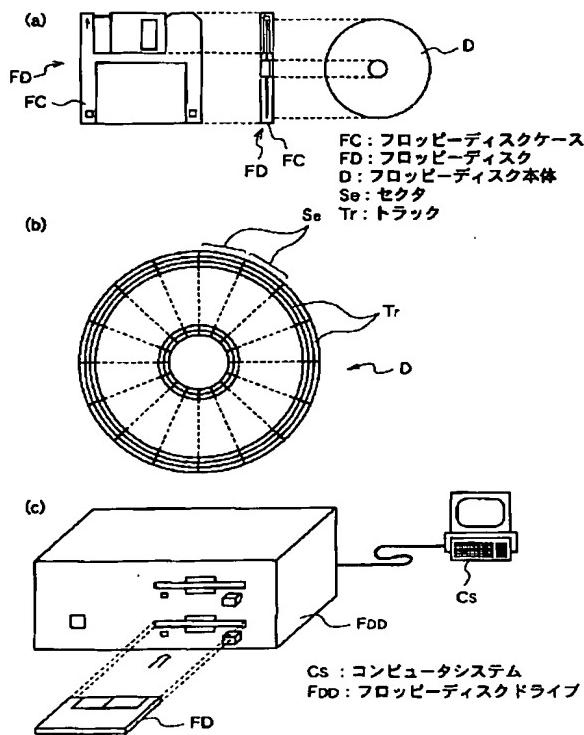
【図 7】



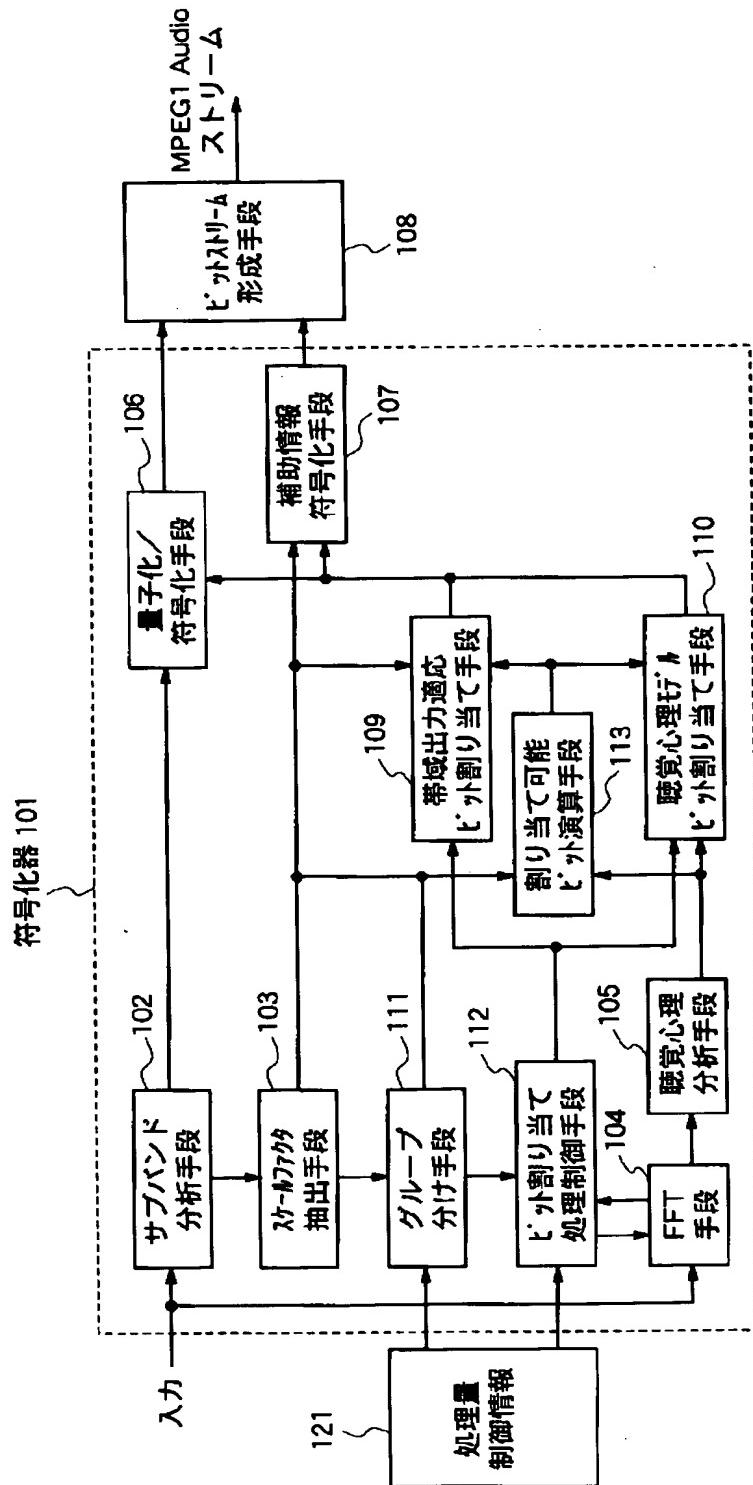
【図 8】



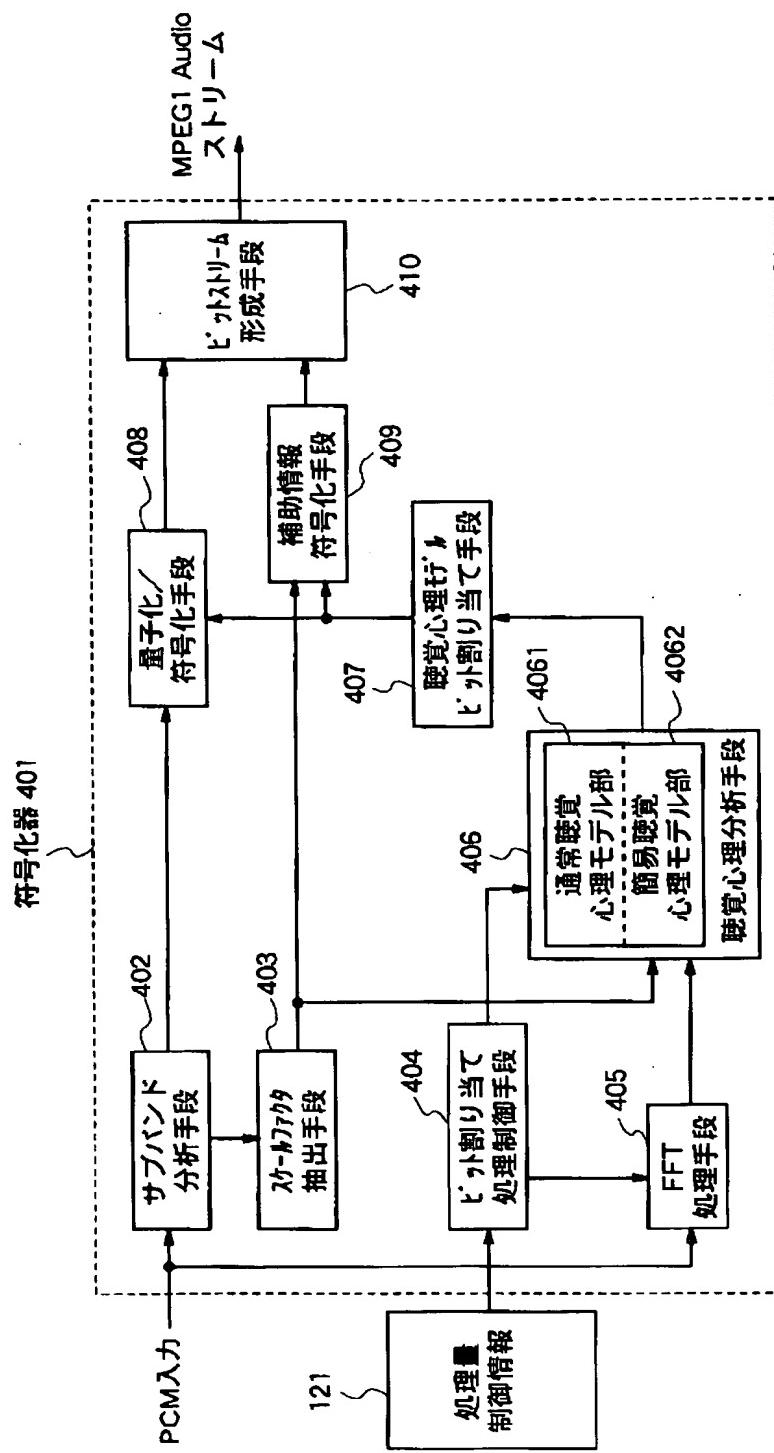
【図 10】



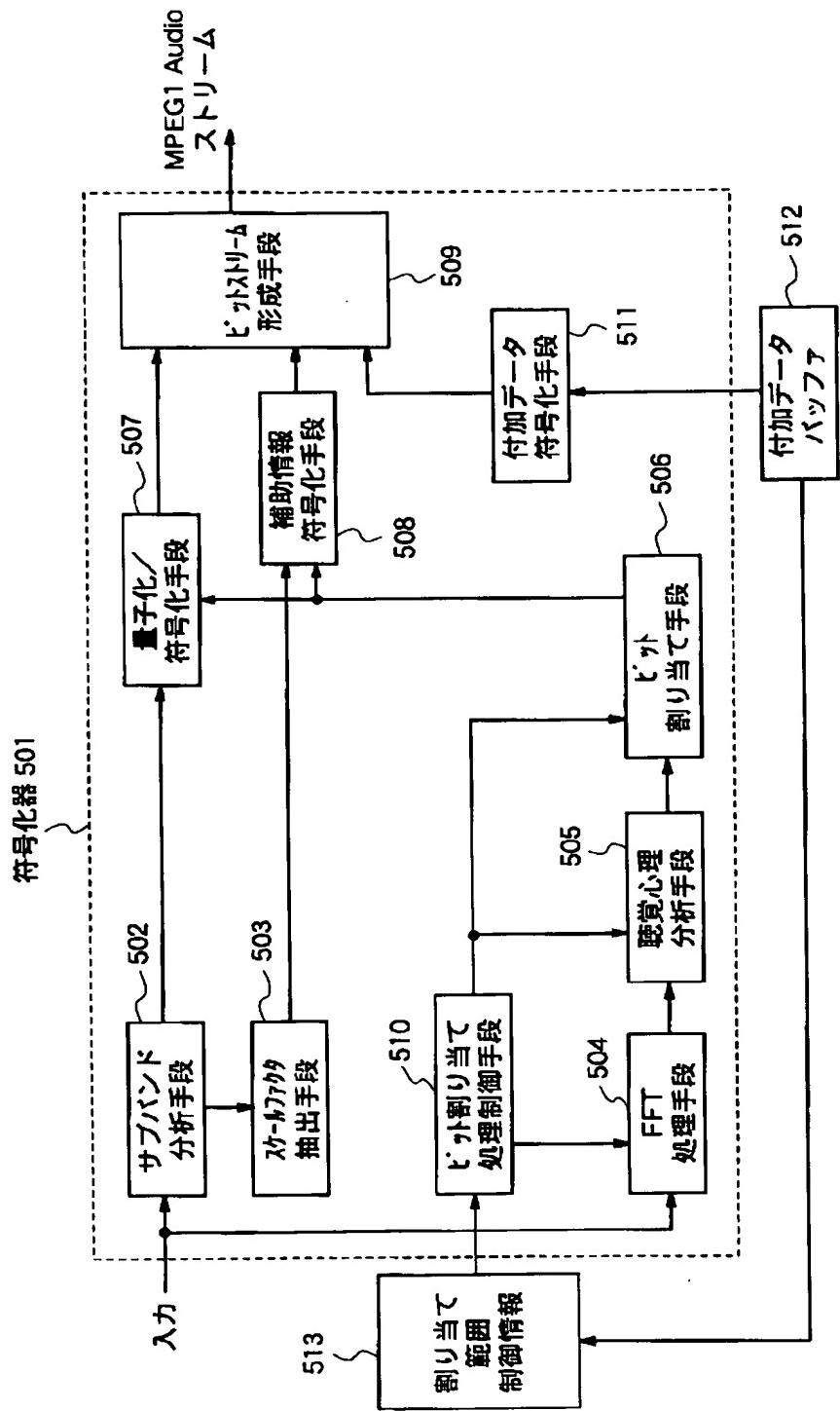
【図9】



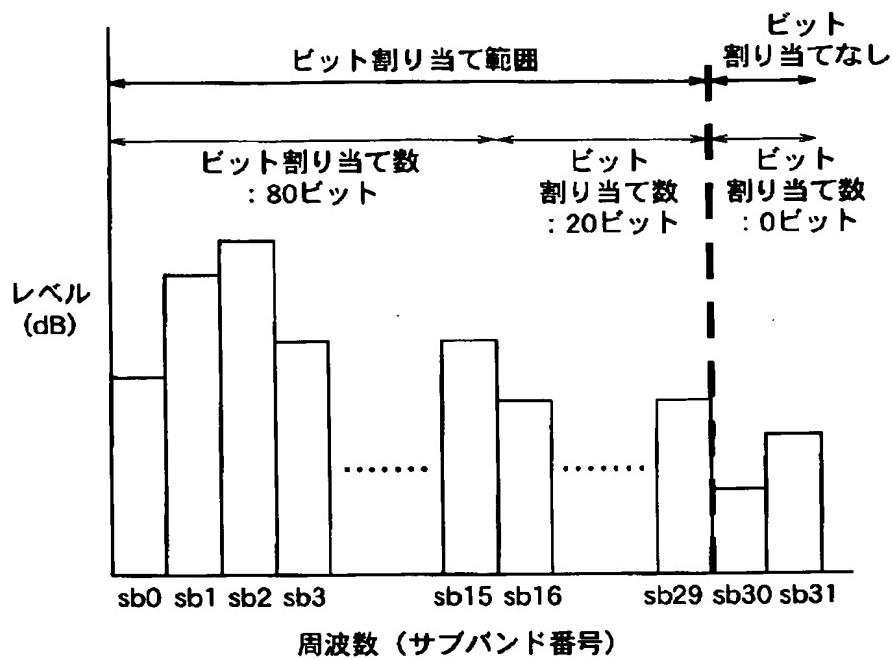
【図1-2】



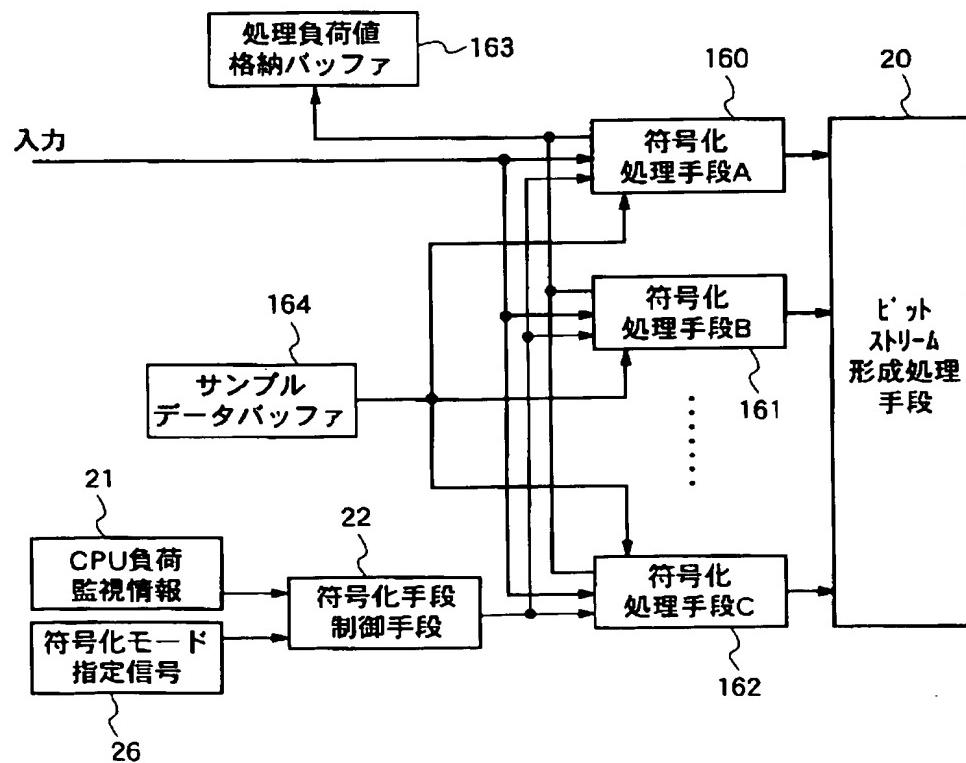
【図14】



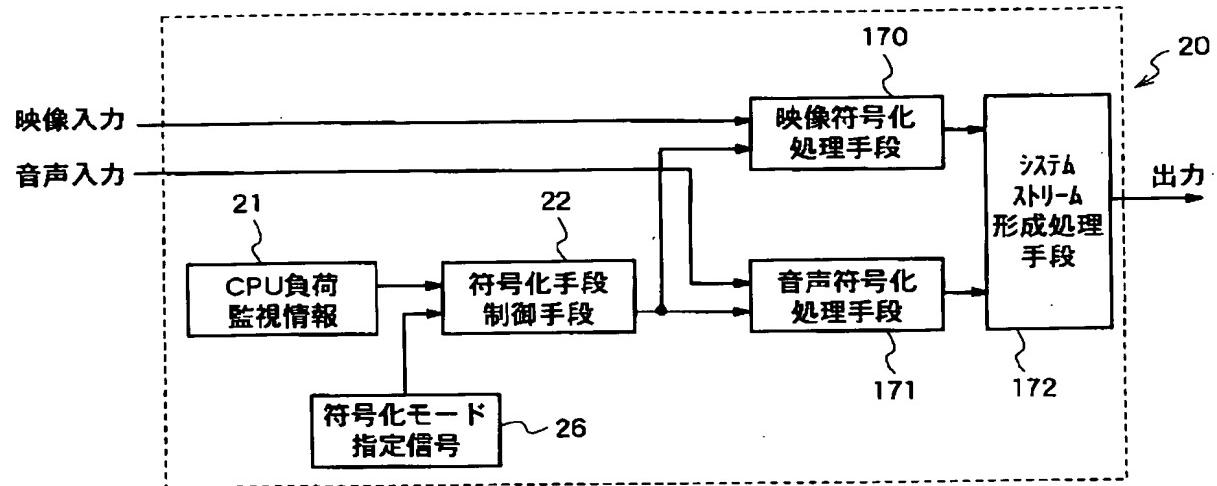
【図 1.5】



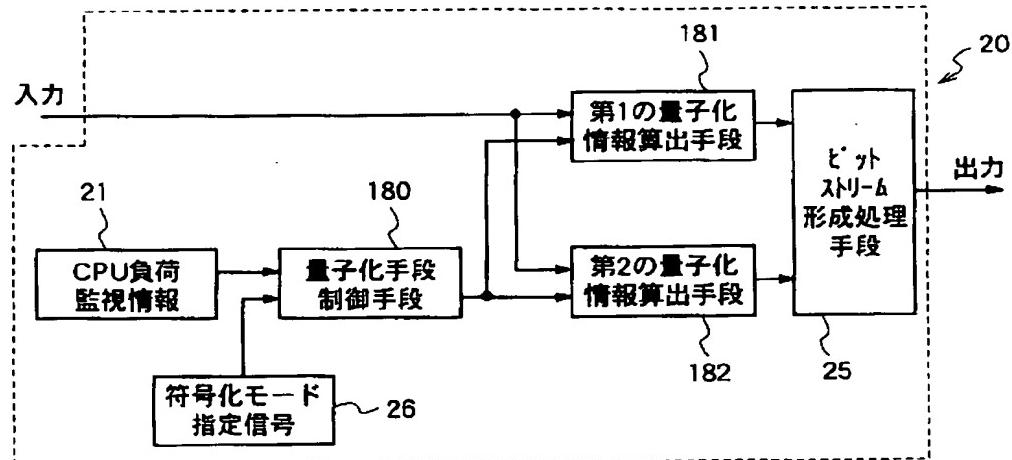
【図 1.6】



【図17】



【図18】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.